(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 November 2004 (18.11.2004)

PCT

(10) International Publication Number
WO 2004/099441 A2

(51) International Patent Classification⁷: C12Q 1/68, C12N 15/10, C07H 21/00

(21) International Application Number: PCT/DK2004/000325

(22) International Filing Date: 6 May 2004 (06.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2003 00711 9 May 2003 (09.05.2003) DK

(71) Applicant (for all designated States except US): HYSCITE DISCOVERY A/S [DK/DK]; Dyckerslundvej 6, 1, DK-5000 Odense C (DK).

(72) Inventors; and
(75) Inventors/Applicants (for US only): MØLLER, Thorleif [DK/DK]; Dyckerslundvej 6, 1, DK-5000 Odense C (DK). PEDERSEN, Kim [DK/DK]; Solfaldsvej 18-11, DK-5000 Odense C (DK).

(74) Agent: PLOUGMANN & VINGTOFT A/S; Sundkrogsgade 9, P.O. Box 831, DK-2100 Copenhagen Ø (DK).

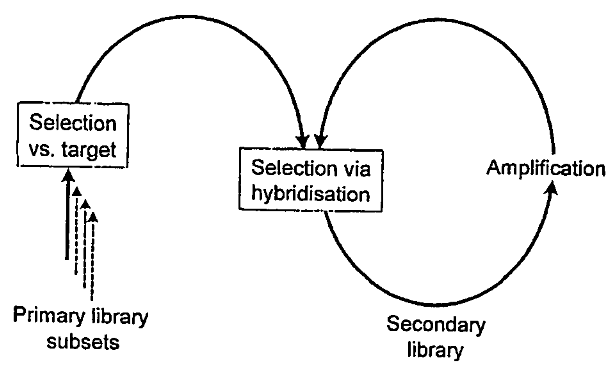
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: SELECTION AND EVOLUTION OF CHEMICAL LIBRARIES



(57) Abstract: The present invention relates to a method for screening libraries of molecules showing specific interaction, such as binding activity or catalytic activity, with a target molecule. The method makes use of a primary library, which comprises the candidate molecules of the library marked with nucleic acid tags and a secondary library, which is used for amplifying and identifying the nucleic acid tags of the molecules in the primary library. The method may e.g. be used for the identification of new pharmaceuticals in which case the target molecule could be a receptor molecule.

WO 2004/099441 A2

BEST AVAILABLE COPY

SELECTION AND EVOLUTION OF CHEMICAL LIBRARIES

FIELD OF THE INVENTION

The present invention relates to a method for screening libraries of molecules showing specific interaction, such as binding activity or catalytic activity, with a target molecule. The method makes use of a primary library, which comprises the candidate molecules of the library marked with nucleic acid tags and a secondary library, which is used for amplifying and identifying the nucleic acid tags of the molecules in the primary library.

BACKGROUND

There is a widespread interest in efficient screening of large numbers of compounds to identify candidate compounds with a given desired activity. In particular, the pharmaceutical industry invests massive efforts into the screening of large libraries of potential drug compounds to find compounds that affect the activity of pharmaceutically relevant targets. Screened compounds include both natural and synthetic compounds. Natural compounds originate from plants, microorganisms or other sources. Synthetic compounds are the result of tedious, organic chemical synthesis. Either way, it is not trivial to build large collections of compounds.

Traditionally, libraries are screened in physically separate assays, which mean that there are limitations as to the number of compounds that can be tested within reasonable time and cost limits, even using automated high throughput screens. It is evident that performing e.g. 1 million assays is a cumbersome task that requires numerous manipulations. To rationalise the screening process, assay volumes are reduced to a minimum with the risk of jeopardising the robustness of the process.

Aiming to reduce the number of manipulations in the generation and screening of libraries, there has been great interest in the synthesis and screening of mixtures of compounds and within the last decade, a relatively simple way to generate very large libraries has been developed. Thus, using combinatorial chemistry, i.e. by synthesising all possible combinations of a set of smaller chemical structures, one-pot libraries of vast size can be

generated. However, the screening of these large combinatorial libraries is perhaps a bigger challenge than their synthesis. Several approaches have been described.

Lam et al. disclose a split-mix combinatorial synthesis of peptides on resin beads and tested the beads against labelled acceptor molecules. Beads binding acceptor molecules were found by visual inspection, physically removed, and the identity of the active peptide was determined by direct sequence analysis.

Houghten et al. used an iterative selection and synthesis process for the screening of combinatorial peptide libraries. Hexapeptide libraries were used to synthesise 324 separate libraries, each with the first two positions fixed with one of 18 natural amino acids and the remaining 4 positions occupied by all possible combinations of 20 natural amino acids. The 324 libraries were then tested for activity to determine the optimal amino acids in the first two positions. To define the optimal third position, another 20 libraries were synthesised by varying the third position and tested for activity. Using this iterative process of synthesis and selection, an active hexapeptide was identified from a library with a total size of more than 34 million hexapeptides. However, the identified peptide is not necessarily the most active peptide in the library, since the first selection is done on the basis of average activity (and not the presence of 1 or a few good peptides) in the 324 libraries that each contains 160.000 (20^4) different peptides and likewise for the subsequent selections.

Another screening approach is based on genetic methods. The advantage of the genetic methods is that libraries can be evolved through iterated cycles of diversification (mutation), selection and amplification as illustrated in Figure 1A. Hence, the initial library needs only contain very tiny amounts of the individual library members, which in turn allow very large numbers of different library species, i.e. very large libraries. Moreover, the structure of active compounds can be decoded with little effort by DNA sequencing. The power of genetic methods for the screening of large libraries is now generally appreciated and has on numerous occasions been used to find new ligands. The major limitation is that only biological molecules can be screened, i.e. peptides that can be synthesised by the translational apparatus or oligonucleotides that can be copied by polymerases. Therefore various approaches have been suggested for the application of genetic screening methods for libraries composed of non-biological molecules.

Liu et al. have suggested using DNA-templated synthesis as a means of evolving non-natural small molecules, and they are developing methods that can translate the amplifiable information in DNA into synthetic molecules (US 20030113738). Likewise WO 02/103008 describes methods to translate information in DNA into synthetic molecules

An early attempt to combine the genetic screening methods with chemically synthesised molecules was put forward in WO 93/20242 by Lerner et al. They performed two alternating parallel syntheses such that a DNA tag is chemically linked to the structure being synthesised. In their method, each chemical step is encoded by the addition of an identifier codon, which means that individual steps of the synthesis can be decoded by sequencing the DNA tag. Using a split-mix protocol, a one-pot library of two-piece bifunctional molecules can be build. However, a library of this type is not evolvable in the traditional sense because the tag does not specify the synthesis of the compounds, rather the tag only reports the synthesis.

However, in WO 93/20242 it is suggested that affinity selected library members have their retrogenetic tag amplified by PCR. DNA strands that are amplified can then be used to enrich for a subset of the library by hybridization with matching tags. The enriched library subset may then be affinity selected against the target and retrogenic tags again PCR amplified for another round of enrichment of a subset of the library. In this method the number of active library members does not increase during the rounds, because active library molecules cannot be amplified/synthesised by way of their tags. Instead it is attempted to remove the non-specific binders from the library as the process proceeds. For very large libraries, though, the amounts of active library members are very tiny, and extra manipulations needed to enrich a library subset before affinity selection seems unfavourable.

SUMMARY OF THE INVENTION

It is an object of preferred embodiments of the present invention to provide a screening method for libraries, e.g. chemical and biological libraries, said libraries comprising potential candidate molecules having non-amplifiable DNA-tags, having amplifiable DNA-tags or other tags of nucleotide-analogues.

It is another object of the present invention to provide an efficient screening method for screening very large libraries, i.e. libraries with a very high number of potential candidate molecules.

It is yet another object of the present invention to provide an efficient screening method for screening libraries having a high level of compounds with very low or no activity.

It is a further object of the present invention to provide a cost- and time-efficient screening method for smaller libraries.

The present invention relates to methods of screening of libraries using an information transfer to an evolvable secondary library as schematically illustrated in Figure 1B.

The method comprises the steps of

a) providing a secondary library comprising a plurality of Y-molecule species, each Y-molecule species comprising a specific tag species (Y-tag species),

b) providing a primary library comprising a plurality of tagged X-molecule species, wherein the tagged X-molecule species comprises an X-molecule species and a specific tag species (X-tag species), and wherein at least one X-tag species of the primary library is capable of hybridising to at least one Y-tag species of the secondary library,

c) contacting the target molecule with at least a subset of the primary library,

d) selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule,

e) optionally, contacting the secondary library with the X-tag species of the selected tagged X-molecule species,

f) selecting Y-molecule species from the secondary library that are capable of hybridising with an X-tag species of a selected tagged X-molecule species of step d) or are capable of hybridising with the complementary sequence of the X-tag species of a selected tagged X-molecule species of step d),

g) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

h) optionally, repeating steps a), f) and g), wherein the secondary library provided in step a) is derived from a secondary library produced in a previous step g),

i) identifying Y-molecule species of high prevalence in a generation of the secondary library, and

j) identifying, from the primary library, X-molecule species corresponding to the Y-tag species of the Y-molecule species of high prevalence.

In an illustrative example of the present invention, the method may be used for screening potential drug candidates for binding activity against a certain receptor. Here the target molecule could be the receptor e.g. immobilised to a solid phase, the tagged X-molecule species could be 10^9 peptides, each peptide carrying a specific DNA tag species, and the Y-molecule species could comprise DNA tag species complementary to the DNA tag species of the peptide and further carry one or more fixed regions which may be used as binding sites for PCR primers in step g) as mentioned above. The specific interaction between tagged X-molecule species and target molecules might in this case be binding. The peptides of the primary library that bind to the receptor molecules are selected in step d), and their corresponding Y-molecule species are selected in step f) by selecting Y-molecule species that are capable of hybridising to the DNA-tag species attached to the selected peptides.

The selected Y-molecule species may be used for preparing a new secondary library, which will be enriched relatively with respect to Y-molecule species that correspond to peptides that bind well to the receptor. The new secondary library may be used in the next repetition of the steps a)-g) and because it is already selectively enriched, the Y-molecule species of the good binders will hybridise even more efficiently than in the first repetition. The concentration of the Y-molecule species corresponding to X-molecules that are poor binders will be reduced as the repetitions progress with new secondary libraries for each repetition and therefore the Y-molecule species of poor binders will hybridise more inefficiently for each repetition. The steps a)-g) are repeated a number of times and for every repetition, the secondary library is further enriched with respect to the Y-molecule species corresponding to the good binders. Finally, the latest secondary library may be analysed and the Y-molecule species of highest concentration are identified along with their corresponding peptides. The identified peptides may now be studied further in more complex models such as cellular or animal models.

Besides for identifying new drug candidates, the present methods may be used for identifying new enzymes for both industrial and therapeutic use, new antibodies and aptamers e.g. for diagnostics, new catalysts, and so forth.

BRIEF DESCRIPTION OF THE FIGURES

In the following, embodiments of the present inventions will be described with reference to the figures, wherein

Figure 1A shows the principle of the genetic screening methods,

Figure 1B shows the principle of double selection and evolution,

Figures 2A-2D illustrate schematically embodiments of a tagged X-molecule species,

Figures 3A and 3B illustrate schematically embodiments of a Y-molecule species,

Figures 4A and 4B illustrate the steps of the method described in Example 1,

Figures 5A, 5B and 5C illustrate the steps of the method described in Example 2,

Figures 6A, 6B and 6C illustrate the steps of the method described in Example 3,

Figures 7A, 7B and 7C illustrate the steps of the method described in Example 4,

Figures 8A and 8B illustrate the steps of the method described in Example 5,

Figures 9A, 9B and 9C illustrate the steps of the method described in Example 6,

Figures 10A, 10B and 10C illustrate the steps of the method described in Example 7,

Figures 11A, 11B and 11C illustrate the steps of the method described in Example 8,

Figure 12 shows a schematic drawing of a tagged X-molecule species having a small peptide as X-molecule species,

Figures 13A and 13B illustrate the steps of the method described in Example 9, and

Figures 14, 15, 16 and 17 shows results from Example 9.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a method of selecting and/or identifying, among a plurality of molecules, a molecule that is capable of specifically interacting with a target molecule. The method comprises the steps of

a) providing a secondary library comprising a plurality of Y-molecule species, each Y-molecule species comprising a specific tag species (Y-tag species),

b) providing a primary library comprising a plurality of tagged X-molecule species, wherein a tagged X-molecule species comprises an X-molecule species and a specific tag species (X-tag species), and wherein at least one X-tag species of the primary library is capable of hybridising to at least one Y-tag species of the secondary library,

c) contacting the target molecule with at least a subset of the primary library,

d) selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule,

e) optionally, contacting the secondary library with the X-tag species of the selected tagged X-molecule species,

f) selecting Y-molecule species from the secondary library that are capable of hybridising with an X-tag species of a selected tagged X-molecule species of step d) or are capable of hybridising with the complementary sequence of an X-tag species of a selected tagged X-molecule species of step d),

g) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

h) optionally, repeating steps a) , f) and g), wherein the secondary library provided in step a) is derived from a secondary library produced in a previous step g),

i) optionally, identifying Y-molecule species of high prevalence in a generation of the secondary library, and

j) identifying, from the primary library, X-molecule species corresponding to the Y-tag species of the Y-molecule species of high prevalence.

Even though it is preferred, the steps of the screening method need not be performed in exact same sequence as written above. However, it is preferred that step a) and step b) are performed before steps c) - i). Step a) may be performed before step b) or step b) may be performed before step a).

Step e) and f) may be performed before step c) and d), such that Y-tag species are hybridised to X-tag species, before tagged X-molecule species are selected against the target molecule.

Step d) and f) may be performed simultaneously. For example, steps c) to g) may be substituted by steps c-1) to f-1):

c-1) hybridising Y-molecule species of the secondary library with X-tag species of the primary library

d-1) contacting the target molecule with at least a subset of the primary library hybridised to the secondary library

e-1) selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule, thereby also selecting Y-tags hybridised to selected X-tags

f-1) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

In a preferred embodiment of the present invention, each X-tag species of at least 50% of the X-tag species of the primary library, such as at least 60%, 70%, 80%, 90%, 95% or 99%, such as at least 100% of the X-tag species of the primary library are capable of hybridising to at most 20 different Y-tag species of the secondary library such as at most 15, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, or 2 Y-tag species, such as at most 1 Y-tag species. For example, each X-tag species of at least 95% of the X-tag species of the primary library may be capable of hybridising to at most 5 different Y-tag species.

In an embodiment of the present invention, the Y-tag of a Y-molecule species may hybridise to only one tagged X-molecule species of the primary library.

In another embodiment, the Y-tag of a Y-molecule species may be able to hybridise to at least 2 different tagged X-molecule species, such as at least 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000 or 10.000 such as at least 100.000 different tagged X-molecule species.

A Y-tag of a Y-molecule species may be able to hybridise to several tagged X-molecule species at a time. For example the Y-molecule species may be able to hybridise to at least

1 molecule of a tagged X-molecule species at a time, such as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000 or 10.000 such as at least 100.000 molecules of a tagged X-molecule species at a time.

In a preferred embodiment of the invention, the Y-tag of a Y-molecule species may be able to hybridise to at most 1000 molecules of a tagged X-molecule species at a time, such as at most 100, 50, 20, 10, 9, 8, 7, 6, 5, 4, 3 or 2 such as at most 1 molecule of a tagged X-molecule species at a time.

In a preferred embodiment of the present invention, the X-tag species of a tagged X-molecule species are not homologues of the X-tag species of another tagged X-molecule species. Also, it may be preferred that the X-tags of individual molecules of the same tagged X-molecule species are identical, alternatively that they are homologues. The X-tag of identical X-molecules may also be non-homologies, that is, two different tagged X-molecule species may comprise the same X-molecule but comprise different X-tags.

Step e) is optional, thus in one embodiment of the present invention the step e) is not performed. In an alternative embodiment step e) is performed. Instead of performing step e), one may use intermediate libraries for transferring the information of the selected tagged X-molecule species, and consequently, one of the intermediate libraries may be hybridised to the secondary library as an alternative to hybridising the selected tagged X-molecule species to the secondary library.

Step h) is optional, thus in one embodiment of the present invention the step h) is not performed. Alternatively, step h) is performed. Step h) comprises the repetition of steps a), f), and g), wherein the secondary library provided in step a) is derived from a secondary library produced in a previous step g). Step h) may furthermore comprise the repetition of one or more of the steps b), c), d) and e). For example, step h) may comprise the repetition of steps a)-g). In a preferred embodiment of the present invention, it is the newest secondary library, i.e. the secondary library of the latest step g) that is used in the next repetition as governed by step h).

The number of repetitions in step h), may be at least 1 times, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20, 30 times or such as at least 40 times. The number of repetitions may be from 1-100 repetitions, such as 1-3 repetitions, 3-5 repetitions, 5-10 repetitions, 10-15 repetitions, or 15-25 repetitions, such as 25-100 repetitions.

Step i) is optional, thus in an embodiment of the present invention the Y-molecule species of high prevalence are not directly identified. Alternatively, step i) is performed and the Y-molecule species of high prevalence are identified in a generation of the secondary library.

Preferably, it is the newest secondary library that is analysed and/or identified in step i), i.e. the secondary library of the latest step g).

The primary library provided in step b) may be substantially identical in every repetition, e.g. the primary library provided may be a sample from a larger primary library stock solution or the primary library may be prepared following the same recipe in every repetition. Two primary libraries are considered "substantially identical" if the relative standard deviation, between the two libraries, of the weight percentage of each tagged X-molecule species is at most 10%, such as at most 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1% such as at most 0.01%. Alternatively, the primary library provided in step b) may be different from the initial primary library in at least one of the repetitions, such as in 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 20 of the repetitions.

In an embodiment of the invention, a first primary library and a second primary library are used in different repetitions in step h). The first and second libraries may differ in that the X-tags of the tagged X-molecule species of the first library are complementary to the X-tags of the corresponding tagged X-molecule species of the second primary library.

The advantage of using complementary X-tags with corresponding pairs of tagged X-molecules is that any unwanted activity coming from the X-tag that may interfere with the primary selection of step d) will not be detected, since the X-tag of the first library is unlikely to have the same binding activity as its complimentary counterpart in the second primary library. Therefore, if a tagged X-molecule species is selected in step d) due to unwanted activity of the X-tag when the first primary library is used, it is unlikely that the same tagged X-molecule species will be selected when the second primary library with the complementary X-tags are used.

In an embodiment of the present invention the method may furthermore comprise a step of monitoring the amplification product of step g) at least one time. The purpose of the monitoring is to evaluate whether another repetition should be performed or whether the secondary library is ready for identification. The amplification product may be analysed by standard methods as described in Sambrook et al and Abelson, e.g. by sequencing the amplification product of step g) in bulk or by cloning the amplification product and sequencing the individual clones. If the analysis reveals that the secondary library has been significantly enriched with respect to a Y-molecule species one could consider

interrupting the repetitions and proceeding with steps i) and j). Depending on the actual embodiment and based on the results of the analysis, the skilled person will be able to determine the right conditions to stop repeating steps a)-g).

A subset of the primary library may e.g. mean the entire material primary library or it may mean a fraction of the material of the primary library, said fraction having a composition which is representative for the composition of the primary library. Also, a subset of the primary library may mean a fraction of the material of the primary library, said fraction having a composition, which is only representative for the composition of the primary library with respect to some of the tagged X-molecule species.

The primary library comprises a plurality of tagged X-molecule species, wherein a tagged X-molecule species comprises an X-molecule species and a specific tag species (X-tag species), and wherein at least one X-tag species of the primary library is capable of hybridising to at least one Y-tag species of the secondary library.

The primary library may comprise at least 10^2 tagged X-molecule species, such as at least 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14} such as at least 10^{15} tagged X-molecule species. For example, the primary library may comprise 10^3-10^{18} tagged X-molecule species, 10^3-10^6 tagged X-molecule species, 10^6-10^9 tagged X-molecule species, 10^9-10^{12} tagged X-molecule species, 10^{12}-10^{15} tagged X-molecule species or 10^{15}-10^{18} tagged X-molecule species.

Preferably, at least one molecule of a tagged X-molecule species should be present in the primary library. The concentration of a tagged X-molecule species may be at least 10^{-22} M such as at least 10^{-21} M, 10^{-19} M, 10^{-18} M, 10^{-17} M, 10^{-16} M, 10^{-15} M, 10^{-14} M, 10^{-13} M, 10^{-12} M, 10^{-11} M, 10^{-10} M, 10^{-9} M, 10^{-8} M, 10^{-7} M, 10^{-6} M, 10^{-5} M, 10^{-4} M, such as at least 10^{-3} M. For example, the concentration of a tagged X-molecule species in the primary library may be in the range of 10^{-22} M - 10^{-20} M, 10^{-20} M - 10^{-18} M, 10^{-18} M - 10^{-16} M, 10^{-16} M - 10^{-14} M, 10^{-14} M - 10^{-12} M, 10^{-12} M - 10^{-10} M, 10^{-10} M - 10^{-8} M, 10^{-6} M - 10^{-4} M, or 10^{-4} M - 10^{-3} M.

The concentration of a tagged X-molecule species in the primary library may be at most 100 mM such as at most 10^{-2} M, 10^{-3} M, 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, 10^{-8} M, 10^{-9} M, 10^{-10} M, 10^{-11} M, 10^{-12} M, 10^{-13} M, 10^{-14} M, 10^{-15} M, 10^{-16} M, 10^{-17} M, 10^{-18} M, 10^{-18} M, 10^{-19} M, 10^{-20} M, 10^{-21} M, such as at most 10^{-22} M.

The primary library may be on liquid form and may comprise an aqueous solvent. The primary library may also comprise an organic solvent and it may comprise both an organic and an aqueous phase at the same time. In a preferred embodiment, the weight

percentage of water in the primary library is at least 50%, such as at least 60, 70, 80, 85%, 90%, 95%, 97%, 98%, 99%, 99.5% such as at least 99.9%

The primary library may also be attached to a solid phase such as particle or a microsphere. The particle or the microsphere may comprise a material selected from the group consisting of an organic polymer, a metal, a metal oxide, an alloy, a magnetic materiel, and a combination of these materials. The metal oxide may be a silicon oxide such as quartz or glass. The organic polymer can be selected from the group consisting of polyethylene glycol-polyacrylamide,, poly styrene, poly vinyl chloride, poly vinyl alcohol, polypeptides, poly ethylene, poly propylene and poly methamethacrylate and a combination of these materials. Also, the particle or microsphere may comprise a composite material having one or more segments with a material as described above.

The primary library may further comprise an additive selected from the group consisting of a detergent, such as Tween 20, NP 40, octylphenolpoly(ethyleneglycolether) (Triton X-100), CHAPS, CHAPSO, sodium dodecylsulfate (SDS); a preservative, such as sodium azide ; a pH buffer such as a phosphate buffer, Tris, Mops or a HEPES buffer; a salt such as $MgCl_2$, NaCl, KCl, Na-glutamate or K-glutamate ; a water soluble polymer such as polyethylene glycol (PEG) or polyvinyl alchohol (PVA). Examples of other suitable additives may be found in Sambrook et al or other general text books known to the person skilled in the art.

In one embodiment of the present invention the primary library may be a microarray and the individual spots of the array may be the different tagged X-molecule species.

The secondary library comprises a plurality of Y-molecule species, said Y-molecule species comprising a specific tag species (Y-tag species).

The secondary library may comprise at least 10^1 Y-molecule species, such as at least 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, such as at least 10^{15} Y-molecule species. For example, the secondary library may comprise 10^3-10^{18} Y-molecule species, 10^3-10^6 Y-molecule species, 10^6-10^9 Y-molecule species, 10^9-10^{12} Y-molecule species, 10^{12}-10^{15} Y-molecule species or 10^{15}-10^{18} Y-molecule species.

As the secondary library is enriched for Y-molecule species that correspond to tagged X-molecule species that are capable of interacting specifically with the target molecule, the Y-molecule species corresponding to tagged X-molecule species that do not interact with the target molecule are diluted in the secondary library. The concentration of a Y-molecule species in the secondary library may be at least 10^{-23} M, such as at least 10^{-21} M, 10^{-20} M,

10^{-19} M, 10^{-18} M, 10^{-17} M, 10^{-16} M, 10^{-15} M, 10^{-14} M, 10^{-13} M, 10^{-12} M, 10^{-11} M, 10^{-10} M, 10^{-9} M, 10^{-8} M, 10^{-7} M, 10^{-6} M, 10^{-5} M, 10^{-4} M, such as at least 10^{-3} M.

Also, the concentration of a Y-molecule species may be at most 100 mM such as at most 10^{-2} M, 10^{-3} M, 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, 10^{-8} M, 10^{-9} M, 10^{-10} M, 10^{-11} M, 10^{-12} M, 10^{-13} M, 10^{-14} M, 10^{-15} M, 10^{-16} M, 10^{-17} M, 10^{-18} M, 10^{-18} M, 10^{-19} M, 10^{-20} M, 10^{-21} M, such as at most 10^{-22} M. For example, the concentration of a Y-molecule species in the primary library may be in the range of 10^{-22} M - 10^{-20} M, 10^{-20} M - 10^{-18} M, 10^{-18} M - 10^{-16} M, 10^{-16} M - 10^{-14} M, 10^{-14} M - 10^{-12} M, 10^{-12} M - 10^{-10} M, 10^{-10} M - 10^{-8} M, 10^{-6} M - 10^{-4} M, or in the range of 10^{-4} M - 10^{-3} M.

Increasing the concentration of a Y-molecule species in the secondary library may speed up the hybridisation reaction. In a preferred embodiment the overall concentration of the secondary library may be decreased along with the repetitions.

In a preferred embodiment of the present invention, the secondary library of step a) is derived from X-tag species of selected tagged X-molecule species of a previous step d).

The term "derive" should be interpreted broadly as providing a secondary library with the same or similar information contents as the starting material, said starting material may e.g. be the X-tags of the selected tagged X-molecule species or the amplified Y-molecule species of step g). In the present context the information contents means the ratio or percentage of the concentration or weight of each Y-molecule species relative to the total concentration or total weight of the Y-molecule species. This may be exemplified by considering a mixture of amplified Y-molecule species comprising the three Y-molecule species Y1, Y2 and Y3 having concentrations of 2 nM, 47 nM and 1 nM, respectively. The information content of the mixture of amplified Y-molecule species Y1, Y2, and Y3 would thus be 2:47:1 or if expressed as percentages: 4% of Y1, 94% of Y2 and 2% of Y3. To have a similar information content it is preferred that the percentage of a Y molecule species in the derived secondary library is in the range of 50%-150% of percentage in the starting material, such as 60%-140%, 70%-130%, 80%-120%, 90%-110%, 95%-105%, 97%-103%, 98%-102%, 99%-101%%, 99.5%-100.5%, or 99.9%-100.1% such as in the range of 99.99%-100.01%. In the example above, if the percentage of Y2 should be in the range of 95%-105% of the percentage in the starting material, this means that the percentage of Y2 should be in the range of 89.3% (94%*0,95) and 98.7% (94%*1,05). In a preferred embodiment of the invention, to have a similar information content it is preferred that the molar percentage a Y-molecule species is within 50%-150% of the molar percentage of the Y-molecule species in the starting material.

Deriving may also mean providing a secondary library with the same or similar information contents as the 50% Y-molecule species or X-tags of the selected tagged X-molecule species of highest concentration and/or weight% in the starting material, such as the top 40%, 30%, 20%, 10%, 5%, 3%, 2%, 1%, 0.1%, 0.01%, 0.001%, 0.0001% or 0.00001%, such as the top 0.000001% of the Y-molecule species of highest concentration and/or weight% in the starting material.

In a preferred embodiment of the present invention, a next generation secondary library is derived from the starting material by providing a secondary library which has an information content similar the 0.001% Y-molecule species of highest concentration in the starting material, the starting material being the amplification product of step f).

Deriving may also mean providing a secondary library with the same or similar information contents as at most the 1,000,000 Y-molecule species or X-tags of the selected tagged X-molecule species of highest concentration and/or weight% in the starting material, such as at most the 100,000, 10,000, 1000, 500, 250, 100, 50, 30, 20, 15, or 10, 5, 4, 3, or 2, such as the one Y-molecule species or X-tags of the selected tagged X-molecule species of highest concentration and/or weight% in the starting material.

In a preferred embodiment of the present invention, a next generation secondary library is derived from the starting material by providing a secondary library which has an information content similar to at most the 1000 Y-molecule species of highest concentration in the starting material, the starting material being the amplification product of step f).

Deriving may comprise processes such as amplification, dilution, restriction, ligation, purification of the coding or the anti-coding strands of the PCR-product, a purification by a standard method e.g. as described in Sambrook et al. Also, deriving may comprise analysing the contents of the starting material and e.g. synthesising or mixing a library with the same or similar composition.

The result of the monitoring of the amplification product of step g) may be used for calculating or estimating the optimal dilution of the amplification product to yield the next generation secondary library.

The secondary library may be derived e.g. using a process where X-tags, either from a non-selected or selected primary library are PCR amplified, whereafter anti-coding strands of the resulting PCR-product is purified and used as a secondary library.

The terms "coding strand" or "coding part" should be interpreted as the tag species of an X-tag species. The "anti-coding strand" or "anti-coding part" is a tag species that is either complementary to the coding tag species or complementary to a tag species which is a homologue of the coding tag species. In the embodiment where a first and second primary library and where the X-tags of the tagged X-molecule species of a first library are complementary to the X-tags of the corresponding tagged X-molecule species of a second primary library, the X-tags of the first primary library are defined as the coding strands and X-tags of the second primary library are defined as anti-coding strands.

The secondary library of step a) may for example be provided by a method comprising the following steps

1) providing a library comprising a plurality of tagged X-molecule species, wherein the tagged X-molecule species is provided with an amplifiable tag species (A-tag species), said A-tag species comprises a tag species and at least one primer binding site for amplifying said tag species,

the tagged X-molecule species are characterised by being divided into two sub-libraries of tagged X_1-molecule species and tagged X_2-molecule species, wherein the amplifiable tag species (A_1) of the X_1-molecule species may be different from the amplifiable tag species (A_2) of the X_2-molecule species

2) contacting a target molecule with the sub-library of tagged X_1-molecule species,

3) selecting, from the sub-library of tagged X_1-molecule species, tagged X_1-molecule species that interact specifically with the target molecule,

4) contacting a target molecule with the sub-library of tagged X_2-molecule species,

5) selecting, from the sub-library of tagged X_2-molecule species, tagged X_2-molecule species that interact specifically with the target molecule,

6) amplifying the A_1-tag species from the selected tagged X_1-molecule species thereby obtaining the anti-coding parts of the selected A_1-tag species,

7) amplifying the A_2-tag species from the selected tagged X_2-molecule species thereby obtaining the anti-coding parts of the selected A_2-tag species,

8) purifying the coding part of the selected A_1-tag species and purifying the anti-coding part of the selected A_2-tag species,

9) contacting the coding part of the selected A_1-tag species with the anti-coding part of the selected A_2-tag species (or vice versa) under conditions that allow for stringent hybridisation,

10) selecting the anti-coding A_2-tag species of step 9) that hybridise to selected coding A_1-tag species, and

11) using the selected anti-coding A_2-tag species of step 10) as secondary library.

Preferably, the only difference between a tagged X_1-molecule species and the corresponding tagged X_2-molecule species, is the sequence of the primer binding site; the X-molecule species of the two species are preferably identical.

Alternatively, the X_1-tags may be complementary to the X_2-tags which could be used to prevent identification of tagged X-molecule species having X-tags with an unwanted binding activity, i.e. X-tags that, either alone or in combination with X-molecules, have affinity for the target and/or the solid phase.

Alternatively, steps 8)-11) could be performed by

8) purifying the anti-coding part of the selected A_1-tag species and purifying the coding part of the selected A_2-tag species, and

9) contacting the anti-coding part of the selected A_1-tag species with the coding part of the selected A_2-tag species (or vice versa) under conditions that allow for stringent hybridisation,

10) selecting the anti-coding A_1-tag species of step 9) that hybridise to selected coding A_2-tag species, and

11) using the selected anti-coding A_1-tag species of step 10) as secondary library.

The sub-libraries may be two physically separate solutions or may both be mixed in one solution.

Step 11) of the method for providing a secondary library may furthermore comprise at least one step selected from the groups of steps consisting of

11a) amplifying the selected anti-coding A_2-tag species,

11b) purifying the amplification product and

11c) adjusting the concentration of amplification product, e.g. by dilution or up-concentration.

Step 11) may also comprise one or more of the steps selected from the group consisting of amplification, dilution, restriction, ligation, purification of the coding or the anti-coding strands of the PCR-product, a purification by a standard method e.g. as described in Sambrook et al.

Preferably, the amplification product is only the amplified tag species and their complementary parts and not side products of the amplification process such as primer-dimers.

According to the present invention, the tagged X-molecule species comprises an X-tag species linked to an X-molecule species, said X-tag species comprising a tag species as defined herein. Several embodiments of tagged X-molecule species are schematically illustrated in Figure 2A-2D.

The tagged X-molecule species, which is illustrated schematically in Figure 2A comprises an X-molecule species (2) linked via a linker molecule (4) to an X-tag (3). The X-molecule species may be build of X-groups (16), e.g. the five X-groups E, D, C, B and A, and the X-tag (3) may be build of tag codons (5), such as the five tag codons A', B', C', D', E'. The X-groups may be connected in a linear way as illustrated in Figure 2A. Alternatively X-groups may form branched structures. To obtain a branched X-molecule structure at least one multifunctional X-group, said X-group comprising at least two active groups, said active groups are capable of further reaction.

In Figure 2B the X-molecule species (2) of the tagged X-molecule species (1) is a molecule such as a protein, a peptide, a oligonucleotide, a small molecule, etc., and said X-molecule species (2) is linked to the X-tag (3) via a linker molecule (4).

The X-tag species may be linked to the X-molecule species via a linker molecule or via a direct binding. The bond involved in direct binding or in the linking using a linker molecule

may be of a covalent character or of a non-covalent character. The linker molecules may be selected from the group consisting of a di-aldehyde such as a polyethylene glycol, glutaraldehyde, a polymer such as an oligosacharide, a nucleic acid and a peptide. The linker molecule may comprise at least two active groups, said active groups are capable of further reaction.

According to the present invention, the term "nucleic acid", "nucleic acid sequence" or "nucleic acid molecule" should be interpreted broadly and may for example be an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof. This term includes molecules composed of naturally-occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as molecules having non-naturally occurring nucleobases, sugars and covalent internucleoside (backbone) linkages which function similarly or combinations thereof. Such modified or substituted nucleic acids may be preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target molecule and increased stability in the presence of nucleases and other enzymes, and are in the present context described by the terms "nucleic acid analogues" or " nucleic acid mimics". Preferred examples of nucleic acid mimetics are peptide nucleic acid (PNA-), Locked Nucleic Acid (LNA-) , xylo-LNA-, phosphorothioate-, 2'-methoxy-, 2'-methoxyethoxy-, morpholino- and phosphoramidate- comprising molecules or the like.

The polymer of the linker molecule may comprise at least 2 monomers such as at least 5, 10, 15, 20, 50, 100 such as at least 200 monomers. Also, the polymer of the linker molecule may be at least 5 Å long such as at least 10 Å, 15 Å, 20 Å, 30 Å, 50 Å, such as at least 1000 Å long.

The polymer of the linker molecole may be substantially linear and it may be substantially unbranched or branched.

The linker of the tagged X-molecule species may be solid phase such as particle or a microsphere. The particle or the microsphere may comprise a material selected from the group consisting of an organic polymer, a metal, a metal oxide, an alloy, a magnetic materiel, and a combination of these materials. The metal oxide may be a silicon oxide such as quartz or glass. The organic polymer can be selected from the group consisting of polyethylene glycol-polyacrylamide, poly styrene, poly vinyl chloride, poly vinyl alcohol, polypeptides, poly ethylene, poly propylene and poly methamethacrylate and a combination of these materials. Also, the particle or microsphere may be a composite material having one or more segments with a material as described above.

Tagged X-molecule species may be of any stochiometry, i.e. any ratio between X-molecule and X-tag species. Thus, a tagged X-molecule may comprise at least 2 molecules of an X-tag species, such as at least 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000, 10.000, 100.000 such as at least 1.000.000 molecules of an X-tag species. Likewise, a tagged X-molecule species may comprise at least 2 molecules of an X-molecule species, such as at least 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000, 10.000, 100.000 such as at least 1.000.000 molecules of an X-molecule species.

The tagged X-molecule species may further comprise a capture component. The capture component may comprise a capture component selected from the group consisting of an amino group, carboxylic group, thiol group, oligonucleotide, peptide, biotin, imino biotin, an avidin, a streptavidin, an antibody, and functional derivatives thereof.

The term "functional derivatives" means derivative of the capture components, said derivatives having substantially the same or improved capture component capability as compared to the capabilities of a capture component listed above.

Also, the tagged X-molecule may comprise a release component. The release component may be located in the X-molecule, or between the X-molecule and the linker molecule, or in the linking molecule, or between the linker molecule and the X-tag species, or in the X-tag species, or between the capture component and the X-tag species.

The release component may be selected from the group consisting of a selective cleavage site for an enzyme, a cleavage site for a nucleic acid restriction enzyme, a disulfide bridge, a ribonucleotide, a photocleavable group.

The photocleavable group may be an o-nitrobenzyl linker, such as described in Olejnik et al 1 and in Olejnik et al 2.

In another embodiment of the present invention, X-tag species cannot be replicated by polymerases. In this embodiment, the X-tag species is composed of unnatural or modified nucleotides that cannot be replicated by polymerases, but are capable of specific basepairing. Examples of unnatural nucleotides are LNA (locked nucleic acids), PNA (peptide nucleic acids), TNA (threose nucleic acids), 2'OH methylated RNA, morpholinos, phosphorothioate nucleotides etc.

The use of an X-tag species composed of unnatural nucleotides may be desired to change the hybridization characteristics of the X-tag species, its chemical or biological stability, its solubility or other characteristics.

In still another embodiment, the X-tag species may also be the X-molecule of the tagged X-molecule species. A non-limiting example thereof is shown in Figure 2C, where the X-molecule species 2 and the X-tag 3 is the same part of the tagged X-molecule species 1. Optionally, the X-tag species may not be able to be replicated by polymerases. Examples of nucleotides that cannot be replicated by polymerases are LNA, PNA, 2'OH methylated RNA, morpholinos, phosphorothioate nucleotides. Also backbone-substituted oligonucleotides of the above-mentioned may be employed. Such tagged X-molecule species may be used where one desires to find an oligonucleotide that is not recognized by proteins that have evolved to interact with natural nucleic acids, e.g. it may be desirable that the particular oligonucleotide is not degraded by nucleases. Or the use of non-natural oligonucleotide may also be desired because of specific demands on chemical stability, solubility or other characteristics.

In a preferred embodiment of the present invention, the X-tag species of the tagged X-molecule species are hybridised to nucleic acid molecules (during step C)), said nucleic acid molecules may comprise universal nucleotides and/or a sequence complementary to the X-tag species. Not to be bound by theory, this approach may in some cases this may be advantageous, since doubled stranded nucleic acid's are less likely to have affinity against the target or exhibit non-specific binding activity than single stranded nucleic acid's. A non-limiting illustration of this embodiment is shown in Figure 2D. Here, the X-molecule species (2) is linked to the X-tag (3) via the linker molecule (4). The X-tag (3) is furthermore hybridised to a complementary nucleic acid molecule (22).

In another preferred embodiment of the present invention, the X-tag species comprises a primer binding site for amplifying the X-tag species. An X-tag species comprising a primer binding site is called an A-tag species.

A primer binding site may be a fixed region within an X-tag species or Y-tag species, said fixed region may be substantial identical or homologue for all the different species. Thus, the fixed region may be an oligonucleotide sequence that is present in all X-tag species or in all Y-tag species of a primary or secondary library.

In the present context the term "homologue" of a given nucleic acid molecule capable of hybridising to a given target sequence means a nucleic acid molecule which is capable of hybridising to the same given target sequence at the same or similar conditions as the given nucleic acid molecule. E.g. if the given nucleic acid molecule is a single stranded DNA molecule, a corresponding DNA molecule, RNA molecule, LNA molecule or PNA molecule would be considered homologue if it was capable to hybridise to the complementary

sequence of the single stranded DNA molecule at a temperature in the temperature range 40 degrees C - 95 degrees C, such as 50 degrees C - 80 degrees C, 50 degrees C - 75 degrees C, 55 degrees C - 65 degrees C, or 55 degrees C - 62.5 degrees C, such as 55 degrees C - 60 degrees C.

The tagged X-molecule species may be prepared using a method comprising the steps of

a) providing a linker molecule comprising at least a first functional group and a second functional group, said first functional group is capable of receiving a tag codon group, said second functional group is capable of receiving an X-group

b) adding a new tag codon group to the first functional group, said new tag codon group being capable of receiving a further tag codon group,

c) adding a new X-group to the second functional group, said new X-group being capable of receiving a further X-group.

Step b) and c) may be performed in the same reaction mixture or in separate mixtures. It may be preferred that step b) and/or step c) comprise(s) a solid phase reaction. Alternatively, it may be preferred that step b) and/or step c) comprise(s) a liquid phase reaction. Step b) may be performed before step c) or step c) may be performed before step b).

The first X-group could contain e.g. three reaction sites, each allowing addition of another X-group which may or may not contain further reaction sites (functionalities capable of receiving another X-group).

The resulting tagged X-molecule species may be of the type shown in Figure 2A.

The X-group may comprise at least one component selected from the group consisting of an amino acid, a nucleotide, a monosaccharide, a disaccharide, a carbohydrate, derivatives thereof, dimers, trimers and oligomers thereof and any combinations thereof.

The amino acid may be selected from the group consisting of alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine, a synthetic amino acid, a beta amino acid, a gamma amino acid and a peptoid (N-substituted glycine).

The X-molecule species may comprise a component selected from a group consisting of a peptide, a nucleic acid, a protein, a receptor, a receptor analogue, a polysaccharide, a drug, a hormone, a hormone analogue and an enzyme. They may also be selected from the group consisting of a synthetic molecule and a molecule isolated from nature.

The X-molecule species may have a molar weight of at most 5.000 kD (kiloDalton) such as at most 1.000kD, 500 kD, 400 kD, 300 kD, 200 kD, 100 kD, 50 kD, 25 kD, 10 kD, 2000 D, 1000 D, 500 D, 250 D, 100 D such as at most 50 D. In a preferred embodiment of the present invention, the X-molecule species may have a molar weight in the range of 50-2000 D, such as e.g. 150-1500 D, 200-1300 D, 500-1000 D, 50-500 D, 250-1000 D, 1000-1500 D or 1500-2000 D.

The X-molecule species may have a molar weight of at least 500 D, such as 1000 D, 5 kD, 10 kD, 20 kD, 40 kD, 80 kD, 200 kD, 500 kD, such as at least 1000 kD. Also the X-molecule species may have a molar weight in the range of 500 D - 1000 kD, such as 500 D- 5 kD, 5 kD 1000 kD, 5 kD - 50 kD, 50 kD - 200 kD, 200 kD - 500 kD or 500 kD - 1000 kD.

The X-molecule species may comprise at most 500 monomer building blocks and/or X-groups such as at most 100, 50, 40, 30, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, such as at most 3 monomer building blocks and/or X-groups.

The X-molecule species may comprise at least 1 monomer building blocks and/or X-groups such as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, such as at least 50 monomer building blocks and/or X-groups. In a preferred embodiment, the X-molecule species may comprise 2-100 monomer building blocks and/or X-groups, such as 2-10, 2-20, 2-10, 5-10, 5-20, or 10-50 monomer building blocks and/or X-groups.

The X-molecule species may be stable within the temperature range 0 to 95 degrees C such as within 0 to 10 degrees C, 10 to 20 degrees C, 20 to 30 degrees C, 30 to 40 degrees C, 35 to 38 degrees C, 40 to 50 degrees C, 50 to 60 degrees C, 55 to 65 degrees C, 60 to 70 degrees C, 70 to 80 degrees C, 80 to 90 degrees C, such as within the temperature range 90 to 95 degrees C. In an embodiment, the X-molecule species may survive 1 hour of autoclaving at 120 degrees C.

The tagged X-molecule species and/or the X-molecule species may be produced by combinatorial chemistry, e.g. such as described in WO 93/20242 or in Needels et al, e.g. using the split-pool principle.

Also, tagged X-molecule species may be prepared using a convergent synthesis, i.e. preparing X-tag and X-molecule (purified, synthesised, or other) separately, followed by attachment of the X-tag to the X- molecule.

According to the present invention, the Y-molecule species may comprise a Y-tag species and may be capable of being amplified.

The Y-molecule species may furthermore comprise a binding site for a PCR primer, e.g. located at the 3' end of the Y-molecule species at the 5'end or at both ends.

A schematic illustration of a Y-molecule species is shown in Figure 3A and 3B. In Figure 3A, the Y-tag (11) comprises the five tag codons (5), namely A', B', C', D' and E'. The Y-tag (11) is flanked by a first fixed region (13) and a second fixed region (14). One of the fixed regions (13) or (14) may be used as a primer binding site during a PCR process. Alternatively, as shown in Figure 3B, the Y-tag (11) may comprise only one fixed region (13).

The binding site may either be a part of the tag species or may not be a part of the tag species.

The Y-molecule species may further comprise a capture component selected from the group consisting of an amino group, a carboxylic group, a thiol group, a peptide, an oligonucleotide, a biotin, an avidin, a streptavidin, an antibody, and functional derivatives thereof.

In a preferred embodiment, the capture component is located at the end of the Y-molecule species.

The Y-molecule species may comprise detectable groups such as radiolabelled groups or fluorescent markers.

The Y-molecule species may further comprise a release component. The release component may be located in the Y-tag species, between the capture component and the Y-molecule, or between the Y-tag species and the binding site for the PCR primer, or at the end of the Y-molecule species. The release component may be selected from the group consisting of a selective cleavage site for an enzyme, a cleavage site for a nucleic acid restriction enzyme, a ribonucleotide, and a photocleavable group.

The photocleavable group may be an o-nitrobenzyl linker.

In a preferred embodiment of the present invention, the Y-molecule species are selected so that the Y-molecule species have substantially no intrinsic binding activity or affinity for the tagged X-molecule species and/or the target molecule. Y-molecule species may preferably have affinity against corresponding tagged X-molecule species, but not against target molecule or other tagged X-molecule species. Y-molecule species which may be unsuitable for use in the present method due to a high level of non-specific or intrinsic binding may be identified by screening the Y-molecule species for intrinsic binding.

The target molecule can be any given molecule or structure to which one wishes to find a ligand. Therapeutically relevant target molecules are mostly proteinaceous molecules. The target molecules may be selected from the group consisting of a protein, a hormone, an interleukin receptor, ion channels, a ribonucleoprotein and a prion.

The protein may be selected from an interleukin, an antibody, an enzyme, a membrane protein, a membrane bound protein, an intracellular protein and an extracellular protein.

Moreover, a target molecule need not necessarily be a single protein. Instead, the target molecule may be a complex of several proteins, a cell membrane, a fragment of a cell membrane e.g. having a lipid double layer, or a cell organ, e.g. golgi apparatus, endoplasmatic reticulum, mitochondria, etc, an entire cell, groups of cells or a tissue. In an embodiment of the present invention, it may be desirable to find molecules that are transported into a cell instead of binding to a particular place on or in the cell. In an embodiment, a molecular library may be incubated with target molecule cells for a certain time and molecules that are transported into the cell may be recovered by e.g. phenol extraction of the cells followed by ethanol precipitation.

When dealing with cellular target molecules, it may be preferred that the X-tag species comprise a biotin-group or a similar capture component to facilitate recovery.

The target molecule could also be a nucleic acid such as a RNA molecule (e.g. tRNA, rRNA, mRNA, miRNA etc.) or a given DNA sequence. Also metabolic intermediates, e.g. stabilised intermediates, may be employed as target molecules.

The target molecule could also be a transition-state analogue, e.g. if one wishes to find new catalysts.

The cell may be a eukaryote cell such as a plant cell, a mammalian cell or a yeast cell or the cell may be a prokaryote cell or the cell may be an archae.

Also, the target molecule may be a virus or a fragment of a virus.

In an embodiment, the concentration of the target molecule used in step c) is kept as low as possible to reduce non-specific binding, while at the same time allowing binding and selection of X-molecules binding specifically to the target molecule. E.g. assuming that a 10^9 library with a total concentration of 100 µM is used and tagged X-molecule species with a K_d value for interaction with the target of less than 10^{-9} M are desired, the appropriate concentration of target can be calculated using the law of mass action. The concentration of individual tagged X-molecules in the library is: 100 µM / 10^9 =10^{-13} M and using the law of mass action, one may calculate the target concentration that allows 99% binding:

Target concentration: (10^{-9} M x $0,99\times10^{-13}$ M) / 0,01 x 10^{-13} M = 9,9 x 10^{-8} M. Thus, at a target concentration of app. 1 nM, 99% of tagged X-molecules with a k_d of 10^{-9} M will be bound to the target at equilibrium.

Also, several different target molecule concentrations may be used such that the primary library is first selected against a relatively low target concentration, and then successively selected against increasing concentrations of target. For each target concentration, a separate secondary library is used. In this way, tagged X-molecules may be identified according to their binding affinity (k_d).

For example, in step c) the ratio between the average number of molecules per tagged X-molecule species and the number of target molecules may be at least 1:10^1, 1:10^2, 1:10^3, 1:10^4, 1:10^5, 1:10^6, 1:10^7,1:10^8, 1:10^9, 1:10^{10}, 1:10^{11}, 1:10^{12} or 1:10^{13}, such as at least 1:10^{14}.

Also, in step c) the ratio between the total number of molecules of all tagged X-molecules species and the number of target molecules may be at most 10^{15}:1 such as at most 10^{14}:1, 10^{13}:1, 10^{12}:1,10^{11}:1, 10^{10}:1,10^9:1,10^8:1, 10^7:1, 10^6:1 10^5:1, or 10^4:1, such as at most 10^3:1.

The tag species comprises a sequence of tag codons, said tag codon is capable of binding to a tag codon with a complementary sequence. The binding occurs preferably by hybridisation.

In a preferred embodiment, the tag species are capable of specific Watson-Crick basepairing and replication by polymerases in PCR.

A tag codon may comprise at least one nucleotide, such as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, such as at least 50 nucleotides.

The sequence of tag codons within a tag species may comprise at least 1 tag codons, such as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, such as at least 20 tag codons.

The tag species may be orthogonal meaning that tag codons and tag codon sequences are selected and/or designed so that no tag species can partly of fully hybridise to another tag species within the temperature range 55-70 degrees C. Tag codons may be designed for example by employing methods described in US 5,635,400 (Minimally Cross-Hybridising Sets of Oligonucleotide Tags).

The tag species may be prepared by standard phosphoramidite oligonucleotide synthesis such as described in Sambrook and in Abelson. However, if a hexacodon tagging system is used, i.e. if the codons comprise six nucleotides, it may be desirable to use hexanucleotide phoshoramidites as building blocks, instead of mononucleotide phosporamidites, as this will result in sixfold fewer couplings in the oligonucleotide synthesis. The same applies if employing a pentacodon, heptacodon tagging system or similar systems.

The ratio between the concentration of a tagged X-molecule species in the primary library and the concentration of its corresponding Y-molecule species in the secondary library will vary from application to application and it will furthermore vary during the repetitions of the method.

In the first cycle of the method of the present invention it may be preferred that the ratio between the concentration of a tagged X-molecule species in the primary library and the concentration of its corresponding Y-molecule species in the secondary library at least 1:10^{10}, such as at least 1:10^6, 1:10^5, 1:10^4, 1:10^3, 1:10^2, 1:10^1, 1:1, 10^1:1, 10^2:1, 10^3:1, 10^4:1, 10^5:1, or 10^6:1, such as at least 10^{10}:1.

The specific interaction between the target molecule and the tagged X-molecule species is an important process and many levels and combinations of specific interaction are envisioned.

The specific interaction is an interaction selected from the group consisting of the binding of a tagged X-molecule species to the target molecule, conformational changes of the tagged X-molecule species and/or the target molecule, the binding of an tagged X-molecule species to the target molecule, enzymatic activity from the tagged X-molecule species on the target molecule, enzymatic activity from the target molecule on the tagged X-molecule species, enzymatic activity complex of the tagged X-molecule species and target molecule, effects in cells, tissue and animals mediated by the target molecule upon binding of the tagged X-molecule species, and any combination thereof.

In an embodiment of the present invention, it is only the X-molecule of the tagged X-molecule species that interacts specifically with the target molecule, whereas in another embodiment it is the combination of X-molecule and X-tag species that is responsible for the interaction. One may experience tagged X-molecule species, in which the X-molecule species alone is not able to interact specifically with the target molecule, but where the combined X-molecule and X-tag species is capable of interacting with the target molecule.

According to the present invention, the methods of selection may be any suitable methods known in the art of screening and selection, e.g. as described in Abelson.

When the specific interaction is binding between the tagged X-molecule species and the target molecule, one may use a selection method comprising the steps of

a) contacting the primary library with a target molecule bound to a solid phase

b) allowing the tagged X-molecule species to bind to the solid phase bound target molecules

c) washing away unbound tagged X-molecule species, thereby leaving, bound to the solid phase, only tagged X-molecule species capable of binding to the target molecule,

d) optionally, eluting the tagged X-molecule species capable of binding to the target molecule from the solid phase,

thereby selecting the tagged X-molecule species capable of binding to the target molecule.

Binding conditions can be adjusted such as to minimize unspecific binding of the tagged X-molecule species in the selection process.

The temperature during the selection of tagged X-molecule species capable of interacting specifically with the target molecule is preferably within the range of 0 to 100 degrees C such as within the temperature 0 to 10 degrees C, 10 to 20 degrees C, 20 to 30 degrees C, 30 to 40 degrees C, 35 to 38 degrees C, 40 to 50 degrees C, 50 to 60 degrees C, 55 to 65 degrees C, 60 to 70 degrees C, 70 to 80 degrees C, 80 to 90 degrees C, such as within the temperature range 90 to 100.

The time in which the specific interaction between the tagged X-molecule species and the target molecule occurs may be within the range 0.001 sec- 20 days such as within 0.001-0.01 sec, 0.01-0.1 sec, 0.1-1 sec, 1-30 sec, 30-60 sec, 60 sec to 1 minute, 1 minute - 20 minutes, 20 minutes to 60 minutes, 60 minutes to 5 hours, 5 hours to 12 hours, 12 hours to 1 day, 1 day to 3 days, 3 days to 6 days, such as within 6 days to 20 days.

In an embodiment of the present invention, substantially all target molecules are bound in the same spatial fashion relative to the solid phase surface. In another embodiment, substantially all target molecules present the same parts, such as epitopes, moieties, sequences etc., of the target molecule to the tagged X-molecule species.

The primary library can be contacted to a target molecule in a number of different experimental settings. Most often the target molecule is present in the solid phase and the primary library in the liquid phase. I.e. the target molecule has been immobilised on a solid matrix. Alternatively, the target may be immobilized after contacting the primary library. The target molecule may be immobilized using CNBr activated sepharose or the target molecule may be biotinylated and immobilized on streptavidin sepharose beads or magnetic streptavidin beads (e.g. Dynabeads® M-280 Streptavidin). Also, filterbinding to can be employed, e.g. to nitrocellulose filters. A great variety of methods for immobilisation of target molecules are known to those skilled in the art. The target molecule may also be present in the liquid phase together with the primary library and the primary library may be present in the solid phase with the target molecule being in the liquid phase. The solid phase may be various kinds of beads as mentioned above, but also microchips/arrays and the like can be employed. The liquid phase will most often be aqueous, the exact composition depending on the particular affinity selection. Hence, the pH of the aqueous media can be controlled using buffer systems such as MOPS, Tris, HEPES, phosphate etc, as can the ionic strength by the addition of appropriate salts. Moreover, it may be desirable to include non-polar, polar or ionic detergents such as NP-40, Triton X-100, Chaps, SDS etc.

Various approaches (not related to incubation conditions, i.e. buffer, temperature, etc.) can be used to reduce selection of non-specific binders. E.g. the library may be counter selected against the solid phase without target molecule, before being selected against the solid phase with target molecule. Moreover, specific binders may be specifically co-eluted with the target molecule, e.g. by cleaving the linker (e.g. photocleavage) that attaches the target molecule to the solid phase. Also competitive elution using known ligands of the target may be used or elution with excess soluble target.

The liquid phase is not limited to aqueous media, as organic solvent may also be employed, those being e.g. DMF, THF, acetonitrile, and organic – aqueous mixtures as well as two phase systems.

The binding reaction may be performed at any desired temperature. If the target molecule is e.g a therapeutically relevant human molecule, the binding reaction may be performed at 37 °C. And for target molecules from thermophilic bacteria a higher temperature can be employed, as well as low temperatures for target molecules from psychrophile organisms, not to preclude any temperature for any target molecule.

The time period for incubation of the binding reaction can be from minutes to hours and even days. The incubation can be adjusted such that the binding reaction is at thermodynamical equilibrium. Moreover, it is possible to select for fast binders (large K_{on} value) by incubating a short time. Likewise, it is possible to select for binders with small K_{off} values by washing the binding reaction and selecting primary library members that stay bound after a chosen time period. Additionally, fast on - fast off binders can be selected by the same method of washing and selecting after a chosen (shorter) time period.

In an embodiment of the present invention, it is possible select for various strengths of binding between the target molecule and the tagged X-molecule species by controlling the conditions during the washing and by controlling the number of washing steps. E.g. if 10 washing steps are performed during the selection process the selected tagged X-molecule species may tend to bind more strongly to the target molecule than if only 2 washing steps were performed.

The amount or concentration of target molecule may be identical or different for each selection round. In one particular embodiment, the amount of target is decreased as the process proceeds.

According to the present invention, the selection of Y-molecule species comprises hybridising a Y-molecule species to the X-tag species of a tagged X-molecule species.

The hybridisation is preferably performed at stringent conditions. The skilled person is readily able devise suitable conditions for the hybridization reactions, assisted by textbooks such as Sambrook, Ausubel et al and Anderson.

The selection may comprise a process selected from the group consisting of amplification, extraction, binding to hydroxyapatite, an enzymatic digest and a hybridisation to a strand immobilized on solid phase followed by a washing step.

The secondary library may be hybridized to X-tag species in a number of ways. If the selected X-molecule species have a stable interaction with their target molecules, the secondary library can be hybridized to X-tag species of tagged X-molecule species fixed to their target molecules. After washing away non-binding Y-molecule species (non-hybridized), hybridized Y-molecule species may be eluted by denaturation with high pH, high temperature or other before PCR amplification. However, it can also be feasible to use the entire binding reaction as template in the PCR reaction, i.e. the solid phase is employed directly in the PCR reaction.

Alternatively, selected X-molecule species can be eluted from the target prior to hybridization with the secondary library. Elution may be done by changing the buffer, e.g. changing ionic strength, pH, detergents, etc., or by raising the temperature. If ligands are sought that bind to the same site of the target molecule as another known ligand the latter may be used for competitive elution. The eluted X-molecule species can then be hybridized to the secondary library in solution, in which case the double stranded product may be recovered by hydroxyapatite chromatography. Alternatively, the X-tag species may be provided with a capture component such as biotin to facilitate recovery. In this case, eluted X-molecule species are hybridized to Y-molecule species in solution and hybridized Y-molecule species recovered by binding X-molecule species to streptavidin beads through a biotin capture component. Eluted X-molecule species can also be immobilized before hybridization.

Various factors may be employed to affect the hybridization reaction, e.g. pH, ionic strength, proteins that affect the rate or fidelity of hybridization, temperature and time of incubation. Also quaternary ammonium salts or betaine, that suppress the effect of base composition making melting temperature, T_m, only dependent on the length of hybrids, can be added. Moreover, the addition of detergents has been reported to speed up the rate of hybridisation. Also, the X-tag species itself may be designed to facilitate hybridization by

employing modified or non-natural nucleotides such as PNA, LNA, 2'O-methylated RNA etc. Further, the sequence content of X-tag species may be designed to facilitate the hybridization reaction. In such a tagging system, it will be desirable to minimize cross-hybridization between non-complementary X-tag species and Y-molecule species. If e.g. 10% of a given X-tag species cross-hybridized to non-complementary Y-molecule species, this may not be much of a problem, since the method is iterative and the fittest Y-molecule species will eventually win. However, minimization of cross-hybridisation may be desirable to minimize the time a given X-tag species will spend on sampling Y-tag species before making a productive encounter with its complementary Y-molecule. This may be of importance for very large libraries, where the number of individual molecules is low. After PCR amplification of selected Y-molecule species, the resulting second-generation secondary library is purified using standard methods (spin-column, gel filtration, gel purification or other) and its concentration adjusted before hybridization with another subset of the primary library selected against the target molecule. Most often, only the anti-coding strand of the PCR product is desired for the secondary library, because the coding strand will compete with the X-tag species for hybridisation to anti-coding strands. Therefore, the anti-coding strand may be purified by elution from immobilized coding strands on streptavidin or by purification from PAGE, as described in the Examples.

If it is desired to speed up the hybridization time in the following rounds, the concentration of the secondary library can be adjusted such as to have Y-molecule species corresponding to active X-molecule species in molar excess (e.g. 10, 50 or 100 fold). Otherwise, the molar ratios can be adjusted such as to reflect a 1 to 1 molar ratio. In the first round, the fold of enrichment in the secondary library can be estimated by measuring the part selected using e.g. radiolabelled Y-molecule species.

In a preferred embodiment, the concentration of Y-molecule species in the secondary library may be adjusted by amplification and/or dilution after each round.

Thus in Example 1 as an example, the part of the primary library that does not bind to the solid phase can be pre-hybridised to the Y-molecule species of the secondary library, before the selected tagged X-molecule species are hybridised to the pre-hybridised secondary library. Also, the primary and secondary library may be hybridised before selection against the solid phase.

For a library composed of tagged X-molecule species such as peptides tagged with an X-tag the problem can be solved in a related way. Thus, a photocleavable biotin may be incorporated in the X-tag. When the primary library is selected against the target molecule, the non-binding tagged X-molecule species are collected and hybridised to the

secondary library. After hybridisation, the hybridisation mixture is illuminated to cleave of the biotin group, whereafter the Y-molecule species of the pre-hybridised secondary library are hybridised to selected tagged X-molecule species, that may still be bound to the target molecule or more likely have been eluted using e.g. SDS, urea or high temperature. The biotin group on selected tagged X-molecule species are used as affinity tag to select secondary library members that correspond to active tagged X-molecule species.

The amplification is performed using a technique selected from the group consisting of Polymerase Chain Reaction techniques (PCR), Strand Displacement Amplification (SDA), Ligation-Rolling Circle Amplification (L-RCA) and their combinations/modifications. These methods are well known to the person skilled in the art and are described in Sambrook. The Y-molecule species may be analysed and identified by standard methods as described in Sambrook et al and Abelson, e.g. by sequencing in bulk or by cloning the amplification product and sequencing the individual clones.

The identification of the Y-molecule species of high prevalence may comprise a step selected from the group consisting of identifying the Y-molecule species with the highest concentration, identifying the Y-molecule species with the highest signal in a hybridisation test, identifying one or more or all Y-molecule species with a concentration and/or signal at a certain threshold, identifying one or more or all Y-molecule species with a concentration and/or signal less than a certain threshold, identifying one or more or all Y-molecule species with a concentration and/or signal above a certain threshold and combinations thereof.

In a preferred embodiment of the present invention, the Y-molecule species are identified as the Y-molecule species, which are present in the PCR product at a concentration at or above a certain concentration threshold.

The identification of the Y-molecule species may be performed with a method comprising the steps of

a) isolating the Y-molecule species from a generation of the secondary library, preferably the newest secondary library, by gel filtration, and

b) identifying one or more Y-molecule species by hybridisation, e.g. to a DNA array or identify one or more Y-molecule species by cloning and sequencing of individual clones.

In a preferred embodiment of the present invention, the tagged X-molecule species that interact specifically with the target molecule is identified from the records respective to which X-tag species that correspond to which X-molecule species. The relevant X-tag species may be identified by identifying the Y-molecule species of high prevalence and either calculating, determining and/or looking up their corresponding X-tag species. The records that relate Y-molecule species to X-tag species and X-tag species to X-molecule species may preferably be handled electronically, e.g. in a computer system.

An additional aspect of the present invention relates to the use of the methods described herein for identifying new enzymes for both industrial and therapeutic use, new antibodies and aptamers e.g. for diagnostic and/or therapeutic use, new catalysts, and so forth. In a preferred embodiment the methods are used for identifying pharmaceutically active compound. The use comprises the preparation of a primary library where the X-molecule species of the tagged X-molecule species are molecules to be tested for pharmaceutical or therapeutic activity against a given disease. The target molecule should preferably have an expected or known relation to the disease. Using the methods described herein, X-molecule species being capable of e.g. binding to the target molecule may be identified and these identified X-molecule species are likely to have pharmaceutical or therapeutic activity against the disease.

EXAMPLES

Examples 1-4 are proof of concept experiments where DNA oligonucleotide libraries are screened to demonstrate that the presented invention can be used as a screening method. Examples 5-8 are extensions of Examples 1-4, which outline how libraries composed of other tagged X-molecule species can be screened. Hence, Examples 5-8 should be generally applicable to libraries composed of tagged X-molecule species.

Example 1: Model system using streptavidin as target molecule and a DNA oligonucleotide comprising a biotin group in a DNA oligonucleotide library as primary library

In this Example, a model library comprising 10^6 different DNA oligonucleotide species in equimolar amounts is screened for binding activity against streptavidin immobilized on sepharose. One particular oligonucleotide in the library contains a biotin-group at its 5'end and it is intended to demonstrate that the identity of this particular oligonucleotide can be found using the present invention. The primary library is prepared by mixing a degenerate

oligonucleotide, which has a total diversity of 10^6, with the biotinylated oligonucleotide, such that the latter is present in equimolar amounts with individual sequences of the degenerate oligonucleotide. Thus it is intended to demonstrate that the present invention can be used to find a signal within about 10^6 fold excess noise. In this context, the word "noise" is used to denote X and Y-molecules that we do not expect to have significant affinity toward the target. Strictly speaking, though, we do not know whether any X or Y-molecules have affinity toward the target, since it is well known that oligonucleotides can take up tertiary structures that bind protein targets with high-affinity and selectivity.

It is important to note that the biotin group serves two roles in Example 1 to 4; the role of a specific interaction in the library relative to the target molecule and the role of a capture component used to manipulate DNA-strands.

The steps of Example 1 are illustrated in Figure 4A and 4B. The two Figures are meant to be combined. The primary library comprises a plurality of tagged X-molecule species (1), one of which is the active tagged X-molecule species (6). The active tagged X-molecule species (5) is marked with a large "X" and the inactive tagged X-molecule species are marked with a small "x". In the present Example 1, the active X-molecule species is a biotin group. Where the biotin group is used as an affinity handle (capture group) for manipulation of DNA strands, the biotin group is indicated by "b". Likewise, where streptavidin sepharose (8) adopts the role of the target molecule it is denoted solid phase bound target and where it is used for manipulation of DNA, it is denoted streptavidin sepharose (18).

Step a) Providing the primary library

The primary library is prepared such as to contain about 10^6 different sequences. This is accomplished using redundant positions during DNA synthesis. To achieve a library with 10^6 different sequences, 12 positions with a redundancy of 2 and 6 positions with a redundancy of 3 are employed ($2^{12} \times 3^6 = 3.0 \times 10^6$). Redundancies are described using the ambiguity table from International Union of Biochemistry (http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html):
M=A or C; R=A or G; W=A or T; S=C or G; Y=C or T; K=G or T; V=A or C or G; H=A or C or T; D=A or G or T; B=G or C or T; N=A or G or C or T.

Oligonucleotide pn1 (primary noise) has a total diversity of 3.0×10^6. The redundancy of each position is indicated below the sequence.

```
pn1          5' MRDTAA KYHGAG YRB[illegible]AC RRBTCT RYVATC MYDTCA
Redundancy      223111 223111 223111 223111 223111 223111
```

The active oligonucleotide containing a 5'biotin ps1 (primary signal) to be present in the primary library is synthesised separately with the following sequence

ps1   5'bAGCTAG TCGGAG CGAAAC GGATCC GCTATA ACCTCG

(b= 5' biotin phosphoramidite catalogue-nr. 10-5950-95 from Glen Research). The underlined sequence is a restriction site for BamHI, used to monitor the evolution of the secondary library. Every third position of pn1 has a redundancy that excludes identity with ps1, i.e. the noise oligonucleotide is designed such that no individual sequence has more than 2/3 identities to ps1. This is to mimic a situation where X-tags have been designed such as to minimize cross-hybridisation.

All oligonucleotides are synthesised using standard DNA oligonucleotide synthesis such as described in (Oligonucleotide Synthesis: A Practical Approach, M.J Gait) and can consequently be purchased from commercial suppliers such as DNA technology A/S, Forskerparken/Science Park Aarhus, Gustaw Wieds Vej 10A, DK-8000 Aarhus C, Denmark, www.dna-technology.com

To prepare 100 µl primary library, ps1 (100 µM) is diluted $3x10^4$ times in TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) + 0.01 % Triton X-100 and 1 µl of this dilution added to 99 µl pn1 having a total oligonucleotide concentration of 100 µM.

Step b) Providing the secondary library

Like the primary library, the secondary library is composed of $3x10^6$ different DNA sequences in equimolar amounts synthesised using redundancies during DNA-synthesis. This is schematically illustrated as the Y-molecule species 11 of Figure 4A. For each coding DNA oligonucleotide in the primary library (tagged X-molecule species), there is a complementary anti-coding DNA oligonucleotide in the secondary library (Y-molecule species). Additionally, the secondary library oligonucleotides have fixed regions in both ends to enable PCR amplification. The noise in the secondary library is represented by oligonucleotide sn1 and the signal is represented by oligonucleotide ss1:

sn1:

5'GATGAT AGTAGT TCGTCG TCAC **TGAHRK GATBRY AGAVYY GTTVYR CTCDRM TTAHYK** AGTC ATGATG AGTAGT TGCTGC

ss1:

5'GATGAT AGTAGT TCGTCG TCAC **CGAGGT TATAGC GGATCC GTTTCG CTCCGA CTAGCT** AGTC ATGATG AGTAGT TGCTGC

The sequence in bold is the anti-coding sequence and the flanking sequences are fixed regions for PCR amplification. Again the underlined sequence is the BamHI restriction site.

PCR primer 1 and PCR primer 2 are used for PCR amplification, the latter PCR primer comprises a biotin group and incorporates the biotin-group into the 5'end of the coding strand of the PCR product:

```
PCR-primer 1:     5'GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2:     5'bGCAGCA ACTACT CATCAT GACT
```

To prepare the secondary library, ss1 (100 µM) is diluted $3x10^4$ times in Te-buffer+ 0.01% Triton X-100 and 1 µl of this dilution added to 99 µl sn1 oligonucleotide stock (100 µM). Note that another 100 µl primary library will be prepared for each round of double selection and evolution, whereas the secondary library will only be prepared once.

Step c) Contacting the primary library with the target molecule

The primary library is contacted with streptavidin immobilized on sepharose (Streptavidin Sepharose High Performance, Cat. No. 17-5113-01, Amersham Biosciences, henceforth also denoted the "solid phase" or "solid phase bound target" when adopting the role of the target and "streptavidin sepharose" when used for manipulations of DNA strands.). Six µl solid phase (20 µl 30% suspension) is equilibrated in 1000 µl binding buffer of 6xSSC + 0.01% Triton X-100 (YxSSC means Y*150 mM NaCl and Y*15 mM trisodium citrate pH 7.0., such that e.g. 6xSSC contains 900 mM NaCl and 90 mM trisodium citrate pH 7.0) and is then incubated in an eppendorf tube for 5 minutes at 65°C with mixing, whereafter the sample is centrifuged at 3000 g and the binding buffer disposed. This washing procedure is repeated twice to equilibrate the solid phase for incubation with the library. The primary library (100 µl) is then added 100 µl 2xbinding buffer (12xSSC + 0.02% Triton X-100 + 4 µg/µl tRNA) before being incubated with the solid phase at 65°C for 30 minutes with mixing.

Step d) Selecting tagged X-molecule species that interact with the solid phase.

After incubation, the solid phase is washed twice as described above with 1000 µl binding buffer to select tagged X-molecule species interacting with the solid phase. In Figure 4A this is shown as a complex 9 between the signal tagged X-molecule species 6 and the solid phase with the target molecule 8.

Step e) Hybridising selected tagged X-molecule species to the secondary library

The secondary library (100µl) is added 1 volume 2xhybridisation buffer (12xSSC + 0.02% Triton X-100 + 4 µg/µl tRNA), before being added to the solid phase with bound tagged X-molecule species. Next, the sample is heated to 85 °C for 5 minutes, followed by incubation at 65 °C for 12 hours.

Step f) Selecting Y-molecule species hybridised to selected tagged X-molecule species

Note that in this step, the particular strong interaction between biotin and streptavidin means that the secondary library can be hybridised directly to selected X-tagged molecules bound to the solid phase. Had the interaction been less strong and the target molecule not been stable during the hybridisation reaction, selected tagged X-molecules could have been immobilized on streptavidin sepharose after selection, as described in Example 5.

After hybridisation, the solid phase is washed two times with 1000 µl hybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C. In Figure 4B this is shown as a new complex between the solid phase with the target molecule 8, the signal tagged X-molecule species 6 and the Y-molecule species 11 which has a Y-tag species which is complementary to the X-tag species of the signal tagged X-molecule species.

Step g) Amplifying the selected Y-molecule species

The washed solid phase may be used directly as template in the amplification step. Alternatively, hybridised Y-molecule species are eluted using spin filtration; the solid phase is suspended in 20 µl 100 mM NaOH, and again separated from the liquid phase using a spin column (Quantum Prep Mini Spin Filters, Cat. No. 732-6027, Bio-Rad). After spinfiltration, 18 µl of the eluate is neutralized by addition of 1 volume (18 µl) 100 mM HCl

and 2/9 volume (4 µl) 900 mM Tris-HCl pH 8.5 and the selected secondary library members are ethanol precipitated by addition of 1/10 volume (4 µl) 3 M Na-acetate pH 4.5 and 3 volumes (120 µl) 96% ethanol followed by 30 minutes centrifugation at 4°C and 20.000 g. Then, the supernatant is disposed, the pellet gently washed with 300 µl icecold 70% ethanol and air-dried. The dried precipitate is dissolved in 28 µl H_2O of which 25 µl is aliquoted into 25 standard PCR reactions each containing: 10 µl OptiBuffer, supplied with the enzyme, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 1, 2 µl 20 µM PCR-primer 2 (comprising a biotin group as a capture component), 63 µl H_2O and 1 µl BIO-X-ACT™ (4 units) Short DNA polymerase (Bioline GmbH, Im Technologiepark, TGZ-2, D-14943, Luckenwalde, cat. no: BIO-21064, www.bioline.com). The reaction is cycled 10 times with 94 °C for 30 sec., 55 °C for 30 sec., 72°C for 60 sec followed by 10 minutes extension at 72 °C. After amplification, all reaction mixtures are pooled and the PCR product is purified by standard gel purification from a 4% agarose gel according to manufacturers instructions (QIAEX II Gel Extraction Kit, Cat. No. 20021, Qiagen, USA, www.qiagen.com). 400 µl H_2O is used to elute the PCR product from Qiaex II beads.

Step h) Preparation of the next generation secondary library

Only the anti-coding strand of the second-generation secondary library is desired for the next hybridisation reaction. Therefore, the PCR product from above is added 1 volume 2x binding buffer and immobilised on 40 µl pre-equilibrated streptavidin sepharose (in Figure 4B it is the streptavidin sepharose 18) by way of the 5'biotin capture component that was incorporated into the coding strand by PCR primer 2. The immobilized PCR product is washed with 1000 µl binding buffer, whereafter the anti-coding strand is eluted with 100 µl 100 mM NaOH as described above using spinfiltration. The eluate is then neutralized, ethanol precipitated and redissolved in 1xhybridization buffer. The concentration of the second-generation secondary library is estimated by UV-absorption and adjusted to a suitable concentration, preferably 10-50 fold lower than the previous generation secondary library (depends on the achieved enrichment).

This second generation library is now ready for next round, where another subset of primary library is selected against the solid phase bound target and selected primary library members hybridized to the second-generation secondary library.

Step i) Repetitions

When a 1000 fold enrichment is achieved in the first round, a total of 10 pmol (i.e. 1000 fold less than in the first round) second generation secondary library can be employed in the next round, in which case ss1 will have the same concentration in the first and second generation secondary library.

Likewise, in the following rounds, a successively lower total concentration of the secondary library can be employed because it evolves to contain a larger fraction of ss1.

The amount of secondary library can also be adjusted to have ss1 in moderate excess (5 - 50 fold) over ps1 for the hybridisation reaction. This provides a safety margin securing information transfer, as well as increasing the rate of hybridisation. If the amount of secondary library is adjusted such as to have ss1 in excess, hybridization times can be adjusted accordingly.

When the total concentration of the secondary library is decreased successively during selection rounds, carrier nucleic acids (e.g. 2 μg/μl tRNA) are added to later generations of the secondary library.

Moreover, the number of cycles in the PCR reactions can be adjusted in later rounds, because the number of secondary library members selected will gradually decrease. The reason for this is that a smaller amount of secondary library is employed for hybridization resulting in less non-specific binding to the solid phase and less specific hybridisation to non-specific tagged X-molecule species.

Step j) Monitoring the evolution of the secondary library

Approximately 0.2 ug (3-4 pmol) of the double stranded secondary library is digested with BamHI to monitor its fraction of oligonucleotide ss1. Digestion is performed with 20 units BamHI in reaction buffer supplied with the enzyme (New England Biolabs, www.neb.com) with incubation for 60 minutes at 37°C. The digested secondary library is resolved on a 4% agarose gel using 1xTBE (0.089 M Tris, 0.089 M Boric acid and 0.002 M EDTA) as running buffer. The fraction of ss1 is estimated by comparing full-length fragments with the fragments resulting from digestion.

Moreover, a fraction of the double stranded secondary library is bulk-sequenced by standard techniques such as described in Sambrook et al. By comparing the sequence of the first generation secondary library with the sequence of later generations of the

secondary library, it can be seen whether the sequence pool is still completely random or whether it has evolved as compared to the starting pool.

Step k) Identifying molecules of high prevalence

A fraction of the double stranded secondary library can be further PCR amplified with cloning primers 1 (5' GCAG *CTCGAG* GATGAT AGTAGT TCGTCG TCAC) and 2 (5' GCAG *CTGCAG* GCAGCA ACTACT CATCAT GACT), which allows directed cloning of the PCR product into pLITMUS™28i (New England Biolabs, #N3528S) using PstI and XhoI restriction sites. After cloning, the identities of a number, e.g. 100, individual clones are determined by sequencing (Litmus forward sequencing primer S1250S, Litmus reverse sequencing primer, S1251S, New England Biolabs), which indicates the composition of the secondary library of the given generation. If all sequenced clones, e.g. 100 clones, are different, more clones may be sequenced, but preferably, the selection process should be continued.

(In the present Example, the Y-molecule species of high prevalence are the three Y-molecule species whose sequences occur the most among the sequenced clones.)

Step l) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

(The links between Y-molecule species and the X-tag species and between the X-tag species and the tagged X-molecule species can be stored in a database of a computer. The Y-tag sequence of the Y-molecule species of high prevalence are submitted to the computer, the computer tracks the relevant relationships in the database and the corresponding tagged X-molecule species and X-molecule species are presented on the monitor of the computer.)

Example 2. Alternative method of preparing the secondary library

In this Example, the first generation secondary library is prepared from a first and a second primary library. The tagged X-molecule species of these two libraries comprise A-tags (X-tags comprising at least one fixed region for PCR amplification) and the libraries only differ in that their A-tags contain different fixed regions for PCR amplification. Both the first and the second primary library are separately selected against the solid phase and

A-tags amplified by PCR. A-tags of first primary library is hybridized to A-tags of the second primary library whereafter hybridized and selected A-tags of the latter are amplified by PCR to generate the secondary library. (One advantage of this method is that the concentration of A-tags corresponding to active X-molecules can be increased relatively to A-tags corresponding to inactive X-molecules before hybridisation.) The steps of Example 2 are illustrated in Figure 5A-5C. The three figures should be combined so that Figure 5A and 5B run in parallel and continue in Figure 5C.

Step 1) Providing the primary libraries

Two primary libraries are prepared, each with a diversity of about 10^9. The coding sequence (shown in bold) of the signal oligonucleotides employed are the same as in Example 1, and again the underlined sequence is a restriction site for BamHI, used to monitor the evolution of the secondary library.

ps2:
5'bGCAGCA ACTACT CATCAT GACT **AGCTAG TCGGAG CGAAAC** <u>**GGATCC**</u> **GCTATA ACCTCG** GTGA CGACGA ACTACT ATCATC

ps3:
5'bCAGTAG TAGCCA ACGGCT AGTA **AGCTAG TCGGAG CGAAAC** <u>**GGATCC**</u> **GCTATA ACCTCG** ATCG TTAGAC GCTATC CGAGTA

The coding sequence of the noise oligonucleotides is designed such as to give a total diversity of about 10^9 or more precisely $2^{30} = 1.1 \times 10^9$. As in Example 1, every third position of the noise oligonucleotides has a redundancy that excludes identity with the signal oligonucleotides.

```
Coding sequence:   MRKKMA KYMRMA YRYMMT RRYKYT RYRMKC MYKKYA
Redundancy:        222221 222221 222221 222221 222221 222221
```

With fixed regions for PCR amplification the oligonucleotides become:

pn2:
5'GCAGCA ACTACT CATCAT GACT **MRKKMA KYMRMA YRYMMT RRYKYT RYRMKC MYKKYA** GTGA CGACGA ACTACT ATCATC

pn3:

5'CAGTAG TAGCCA ACGGCT AGTA **MRKKMA KYMRMA YRYMMT RRYKYT RYRMKC MYKKYA** ATCG TTAGAC GCTATC CGAGTA

The following PCR primers are used:

PCR-primer 1:	5'GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2:	5'bGCAGCA ACTACT CATCAT GACT
PCR primer-3:	5' TACTCG GATAGC GTCTAA CGAT
PCR primer-4:	5'bCAGTAG TAGCCA ACGGCT AGTA

Oligonucleotide ps2 (500 µM) is diluted 1.1×10^7 times in TE buffer + 0.01% Triton X-100 and 5 µl of this dilution added to 495 µl of pn2 (500 µM) to give 500 µl of the first primary library (comprising ps2 and pn2) and likewise for the preparation of the psn3 library. The total amount of individual oligonucleotides in the libraries (500 µl) is now $(6 \times 10^{23} \times 5 \times 10^{-4} \times 5 \times 10^{-4}) / 1.1 \times 10^9 = 1.4 \times 10^8$ and their concentrations are 4.67×10^{-13} M.

Before starting the selection process, second-strand synthesis is performed, because dsDNA is less prone to interfere with selection than ssDNA. For second strand synthesis of psn2, PCR-primer 1 is used and for psn3, PCR-primer 3 is used.

The primary library is split into 10 aliquots of 50 µl each, to which the following is added: 100 µl 300 µM downstream primer, 1000 µl Optibuffer, 600 µl 25mM $MgCl_2$, 160 µl 25 mM dNTP, 100 µl (400-units) Bio-X-ACT™ Short DNA polymerase and 8040 µl H_2O. The ten tubes are incubated in a 94°C water bath for 6 minutes, transported to an 84 °C water bath for 6 minutes, next to 74°C for 6 minutes, 64°C for 6 minutes, and 54 °C for 10 minutes. After annealing of the downstream primer, second strand synthesis is performed at 72°C for 60 minutes in a water bath. Finally, the samples are precipitated by addition of 1/10 volume 3 M Na-acetate pH 4.5 and 3 volumes 96% ethanol and incubation for 30 minutes at –20 °C. The samples are then centrifuged 60 min at 10.000g, the supernatant disposed, and the pellet gently washed twice with 1 ml ice-cold 70% ethanol and air-dried. The dry pellets are redissolved in 100 µl binding buffer and all samples are pooled into a primary library of 1000 µl, that is extracted twice with 200 µl phenol, followed by one extraction with 200 µl chloroform, whereafter the primary library is ready for selection.

Step 2) Contacting the primary libraries with the target molecule

The two primary libraries, psn2 and psn3, are separately contacted with pre-equilibrated solid phase bound target as described in Example 1, step c.

Step 3) Selecting tagged X-molecule species that interact with the target molecule

See Example 1, step d

Step 4) Amplifying the selected A-tags

Second strands (anti-coding strands) are eluted from the solid phase with bound tagged X-molecule species, before serving as templates for PCR amplification; the solid phase is resuspended in 60 µl 100 mM NaOH and spinfiltered, whereafter the eluate is neutralised by the addition of 60 µl 100 mM HCl and 15 µl 900 mM Tris-HCl pH 8.5. Subsequently, 126 µl is aliquoted into 63 standard PCR reactions each containing: 10 µl Optibuffer, 16 µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM upstream PCR-primer, 2 µl 20 µM downstream PCR-primer, 61 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 90 sec followed by 10 minutes extension at 72 °C.

For amplification of the psn2 primary library, PCR primers 1 and 2 are employed. Because PCR primer 2 is biotinylated in its 5'end, the resulting PCR product is biotinylated at the 5'end of the coding strand. Likewise, for amplification of psn3, PCR primers 3 and 4 are employed. Similar to PCR primer 2, PCR primer 4 is biotinylated and consequently the resulting PCR product is biotinylated at the 5' end of the coding strand.

Step 5) Providing the secondary library

a) The psn2 PCR products and psn3 PCR products are ethanol precipitated and redissolved in 500 µl H_2O. Next, the samples are extracted twice with 200 µl phenol, and one time with 200 µl chloroform followed by immobilization on 100 µl pre-equilibrated streptavidin sepharose.

b) The anti-coding strand of the psn2 PCR product is batch eluted by adding 400 µl 100 mM NaOH to the streptavidin sepharose followed by centrifugation of the eppendorf tube. After elution, the streptavidin sepharose containing the psn2 coding strand is washed twice with 1000 µl hybridization buffer.

The anti-coding strand of the psn3 PCR product is eluted with 400 µl 100 mM NaOH using spinfiltration. The eluate is subsequently neutralized, whereafter the ssDNA is ethanol precipitated and redissolved in 400 µl binding buffer + 2 µg/µl tRNA.

c) The streptavidin sepharose immobilised coding strands of the psn2 PCR product are now hybridised to complementary anti-coding strands from the psn3 PCR product, which are next added. Hybridisation is performed by heating the sample to 85 °C for 5 minutes, followed by incubation at 65° for 12 hours.

d) After hybridisation, the streptavidin sepharose is washed two times with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) for 5 minutes at 65°C to select hybridised psn3 anti-coding strands (Y-molecule species)

e) Selected psn3 strands are eluted with 400 µl 100 mM NaOH using spinfiltration, whereafter the eluate is neutralized and ethanol precipitated. The dried precipitate is dissolved in 22 µl H_2O of which 20 µl is aliquoted into 10 PCR reactions each containing: 10 µl optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 3, 2 µl 20 µM PCR-primer 4, 62 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.

f) The resulting PCR product is immobilized on 15 µl streptavidin sepharose, wherafter the anti-coding strand is eluted with 40 µl 100 mM NaOH, followed by neutralisation and ethanol precipitation. The air-dried precipitate is dissolved in 20 µl H_2O to produce the first generation secondary library. The concentration of the second-generation secondary library is estimated by UV-absorption and adjusted to a suitable concentration as described in Example 1, step h).

Step 6) Repetitions

The new secondary library may be used as first generation secondary library in Example 1, thus replacing step b) of Example 1. Furthermore, the first primary library of Example 2 may be used as primary library of Example 1, thus replacing step a) of Example 1.

In the next round, the first primary library is again selected against the solid phase and selected A-tags PCR amplified and immobilized on streptavidin sepharose. The anti-coding strands are then eluted and coding strands hybridized to complementary anti-coding Y-molecule species of the first generation secondary library. Hereby selected Y-molecule species are PCR amplified to generate the second-generation secondary library.

As described in Example 1, step i, the secondary library can be increasingly diluted, because is evolves to contain a larger fraction of signal oligo (ss3), i.e. if the first generation secondary library is 10000 fold enriched in signal oligonucleotides, a 10.000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have ss3 in moderate excess (5-50 fold) over ps1 for the hybridisation reaction. Further, the number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step 7) Monitoring the evolution of the secondary library

See Example 1, step j.

Step 8) Identifying molecules of high prevalence

See Example 1, step k

Step 9) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 3

Example 3 is a modification of Example 2, the major difference being the use of photocleavable biotin groups. When the biotin group adopts the role of X-molecule species, the photocleavable linker allows specific elution of selected tagged X-molecule species. When the biotin group serves as an affinity handle (capture group) that allows simple manipulations of DNA strands, the photocleavable linker adds the possibility of eluting DNA strands that have been immobilized on streptavidin sepharose. The steps of Example 3 are illustrated in Figures 6A-6C. The three figures should be combined so that Figure 6A and 6B run in parallel and continue in Figure 6C.

Step 1) Providing the primary libraries

Two primary libraries are prepared, each with a diversity of about 10^9. The signal oligonucleotides employed are the same as in Example 2, except that a photocleavable linker has been inserted between the biotin group and the X-tag species. This combination of photocleavable linker and biotin is abbreviated pcb.

ps2:
5'pcbGCAGCA ACTACT CATCAT GACT **AGCTAG TCGGAG CGAAAC <u>GGATCC</u> GCTATA ACCTCG** GTGA CGACGA ACTACT ATCATC

ps3:
5'pcbCAGTAG TAGCCA ACGGCT AGTA **AGCTAG TCGGAG CGAAAC <u>GGATCC</u> GCTATA ACCTCG** ATCG TTAGAC GCTATC CGAGTA
(pcb= PC biotin phosphoramidite catalogue-no. 10-4950-95 from Glen Research, USA , www.glenresearch.com)

The coding sequence of the noise oligonucleotides is identical the pn2 and 3, Example 2

The following PCR primers are used:

PCR-primer 1:	5'GATGAT AGTAGT TCGTCG TCAC
PCR-primer 3:	5' TACTCG GATAGC GTCTAA CGAT
PCR-primer 5:	5'pcb GCAGCA ACTACT CATCAT GACT
PCR-primer 6:	5' CAGTAG TAGCCA ACGGCT AGTA
PCR-primer 7:	5'pcb TACTCG GATAGC GTCTAA CGAT

Oligonucleotide ps2 (500 µM) is diluted 1.1×10^7 times in 0.01% Triton X-100 and 5 µl of this dilution added to 495 µl of pn2 (500 µM) to give 500 µl of the psn2 primary library and likewise for the preparation of the psn3 library.
The total amount of individual oligonucleotides in the libraries (500 µl) is now $(6\times10^{23} \times 5\times10^{-4} \times 5\times10^{-4}) / 1.1\times10^9 = 1.4\times10^8$ and their concentrations are 4.67×10^{-13} M.

Second-strand synthesis is performed as described in Example 2, step 1.

Step 2) Contacting the primary libraries with the target molecule

The two primary libraries, psn2 and psn3, are separately contacted with pre-equilibrated solid phase bound target as described in Example 1, step c.

Step 3) Selecting tagged X-molecule species that interact with the solid phase.

After incubation, the solid phase is washed twice with 1000 µl binding buffer to select tagged X-molecule species interacting with the solid phase bound target. Moreover, tagged X-molecule species bound specifically are eluted using the photocleavable biotin linker; the solid phase is resuspended in 75 µl binding buffer and placed on a sheet of parafilm whereafter the sample is illuminated for 6 minutes as described (Olejnik J, Krzymanska-Olejnik E, Rothschild KJ. Photocleavable biotin phosphoramidite for 5'-end-labeling, affinity purification and phosphorylation of synthetic oligonucleotides. Nucleic Acids Res 1996 Jan 15;24(2):361-6). The samples are then spinfiltered and the liquid phase collected.

Step 4) Amplifying the selected A-tag species

The liquid phase containing specifically eluted tagged X-molecule species is aliquoted into standard 60 PCR reactions each containing: 10 µl optibuffer buffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM upstream PCR-primer, 2 µl 20 µM downstream PCR-primer 2, 62 µl H_20 and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.
For amplification of the psn2 primary library, PCR primers 1 and 5 are employed. Because PCR primer 5 is biotinylated in its 5'end, the resulting PCR product is biotinylated at the 5'end of the coding strand. Likewise, for amplification of psn3, PCR primers 6 and 7 are employed which biotinylates the resulting PCR product at the 5' end of the anti-coding strand.

Step 5) Providing the secondary library

a) The psn2 PCR products and psn3 PCR products are ethanol precipitated and redissolved in 500 µl binding buffer + 4 µg/µl tRNA. Next, the samples are extracted twice with 200 µl phenol, and one time with 200 µl chloroform followed by immobilization on 100 µl pre-equilibrated streptavidin sepharose.

b) The anti-coding strand of the psn2 PCR product is batch eluted with 400 µl 100 mM NaOH added to the streptavidin sepharose followed by centrigugation of the eppendorf tube. After NaOH elution, the streptavidin sepharose containing the psn2 coding strand is washed twice with 1000 µl hybridization buffer.

Similarly, the coding strand of the psn3 PCR product is eluted with 400 µl 100 mM NaOH, whereafter the streptavidin sepharose is washed twice with 1000 µl binding buffer. The anti-coding strand is then cleaved of the streptavidin sepharose using the photocleavable biotin linker; the streptavidin sepharose is resuspended in 400 µl binding buffer, placed on a sheet of parafilm and illuminated for five minutes at 325 nm, wherafter the sample is spinfiltered and the eluate collected. The streptavidin sepharose is then washed with another 400 µl binding buffer, spinfiltered and the eluate added to the first eluate. The combined eluate is now ethanol precipitated and redissolved in 400 µl hybridisation buffer.

c) The immobilised coding strands of the psn2 PCR product are now hybridised to complementary anti-coding strands from the psn3 PCR product, i.e. 400 µl psn3 anti-coding strands are added to psn2 coding strands immobilised to streptavidin sepharose. Hybridisation is performed by heating the sample to 85 °C for 5 minutes, followed by incubation at 65° for 12 hours.

d) After hybridisation, the streptavidin sepharose is washed two times with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (3xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C to select hybridised psn3 anti-coding strands (Y-molecule species)

e) psn3 strands selected by hybridisation are eluted by photocleavage as described in step b, whereafter the eluate is ethanol precipitated. The dried precipitate is dissolved in 22 µl H_2O of which 20 µl is aliquoted into 10 PCR reactions each containing: 10 µl optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 6, 2 µl 20 µM PCR-primer 7, 62 µl H_20 and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.

f) The resulting PCR product is immobilized on 15 µl streptavidin sepharose, wherafter the coding strand is eluted with 40 µl 100 mM NaOH, the streptavidin sepharose washed twice with 1000 µl hybridisation buffer, whereafter the anti-coding strand is eluted by photocleavage. Subsequently, the eluate is ethanol precipitated and redissolved in 20 µl hybridisation buffer to produce the first generation secondary library.

Step 6) Repetitions

In the next round, the psn2 primary library is again selected against the solid phase bound target, specifically eluted, selected X-tags PCR amplified and immobilized on streptavidin

sepharose. The anti-coding strands are then eluted and coding strands hybridized to complementary anti-coding Y-molecule species of the first generation secondary library. Hereby selected Y-molecule species are PCR amplified to generate the second-generation secondary library.

As described in Example 1, step l, the secondary library can be increasingly diluted, because it evolves to contain a larger fraction of signal oligo (ss3), i.e. if the secondary library is 10000 fold enriched in signal oligonucleotides, a 10.000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have ss3 in moderate excess (5 –50 fold) over ps1 for the hybridisation reaction. Further, number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step 7) Monitoring the evolution of the secondary library

See Example 1, step j.

Step 8) Identifying molecules of high prevalence

See Example 1, step k.

Step 9) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 4

Example 4 is a modification of Example 2, the major difference being that the hybridisation reaction is performed with both the anti-coding and the coding strand in solution, as opposed to Example 2 and 3, where one strand is immobilised during the hybridisation reaction. The steps of Example 4 are illustrated in Figure 7A-C. The three should be combined so that Figure 7A and 7B run in parallel and continue in Figure 7C.

Step 1) Providing the primary libraries

The two primary libraries employed are identical to the libraries of Example 2.

The following PCR primers are used:

PCR-primer 1:	5' GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2:	5'bGCAGCA ACTACT CATCAT GACT
PCR primer-3:	5' TACTCG GATAGC GTCTAA CGAT
PCR primer-4:	5'bCAGTAG TAGCCA ACGGCT AGTA
PCR primer-8:	5' CAGTAG TAGCCA ACGGCT AGTA

The second nucleotide from the 3' end in PCR primer-8 is a ribonucleotide.

Step 2) Contacting the primary libraries with the target molecule

The two primary libraries, psn2 and psn3, are separately contacted with pre-equilibrated solid phase bound target as described in Example 1, step c.

Step 3) Selecting tagged X-molecule species that interact with the target molecule

See Example 1, step d

Step 4) Amplifying the selected A-tag species

Performed as described in Example 2, except that PCR primers 4 and 8 are used for PCR amplification of selected psn3 molecules.

Step 5) Providing the secondary library

a) The psn2 PCR products and psn3 PCR products are ethanol precipitated and redissolved in 500 µl H_2O. Next, the samples are extracted twice with 200 µl phenol, followed by extraction with 200 µl chloroform.

i) The psn2 PCR product is now immobilized on 100 µl pre-equilibrated streptavidin sepharose and the anti-coding strand of the psn2 PCR product

eluted with 400 µl 100 mM NaOH using spinfiltration. Next, the eluate is neutralised, ethanol precipitated and redissolved in 10 µl hybridisation buffer.

ii) The psn3 PCR product is added 1/10 volume 1 M NaOH and incubated at 80 °C for 5 minutes, which cleaves the anti-coding strand at the ribonucleotide residue in PCR primer-8. Next, the sample is neutralised, ethanol precipitated and redissolved in 500 µl formamide loading buffer. The sample is now heated to 94° for 3 minutes and loaded on a 6% denaturing (8 M urea) polyacrylamide gel and the fragments are resolved until the coding strand has reached approximately the middle of the gel. Hereafter, the exact positions of fragments are determined by UV-shadowing and the gel-piece containing the coding strand is cut out for subsequent passive elution. After elution, the coding strand is extracted with phenol and chloroform and then ethanol precipitated and redissolved in 10 µl hybridisation buffer.

b) The coding strand of psn3 and the anti-coding strand of psn2 are now mixed for hybridisation in solution. Hybridisation is performed by heating the sample to 85 °C for 5 minutes, followed by incubation at 65° for 12 hours.

c) After hybridisation, the volume is increased to 100 µl by addition of binding buffer + 2 µg/µl tRNA, whereafter the sample is added to 5 µl pre-equilibrated streptavidin sepharose and incubated for 30 minutes at 55 °C with mixing.

d) After immobilisation, the streptavidin sepharose is washed two times with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C to select psn2 strands hybridised to psn3 strands.

e) Selected psn3 strands are eluted with 400 µl 100 mM NaOH using spinfiltration, whereafter the eluate is neutralized and ethanol precipitated. The dried precipitate is dissolved in 22 µl H_2O of which 20 µl is aliquoted into 10 PCR reactions each containing: 10 µl optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 1, 2 µl 20 µM PCR-primer 2, 62 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.

f) The resulting PCR product is immobilized on 15 µl streptavidin sepharose, wherafter the anti-coding strand is eluted with 40 µl 100 mM NaOH, followed by neutralisation and ethanol precipitation. The airdried precipitate is dissolved in 20 µl H_2O to produce the first generation secondary library.

Step 6) Repetitions

In the next round, the psn3 primary library is again selected against the solid phase bound target and selected X-tags PCR amplified. The anti-coding strand from the resulting PCR product is hydrolysed with NaOH and the coding strand purified from PAGE. Purified psn3 coding strands are hybridized to complementary anti-coding Y-molecule species of the first generation secondary library in solution, where after hybridised Y-molecule species (psn2 strands) are selected on streptavidin sepharose. Hereby selected Y-molecule species are PCR amplified to generate the second-generation secondary library.

As described in Example 1, step l, the secondary library can be increasingly diluted, because is evolves to contain a larger fraction of signal oligo (ss3), i.e. if the secondary library is 10000 fold enriched in signal oligonucleotides, a 10.000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have ss3 in moderate excess (10 –100 fold) over ps1 for the hybridisation reaction. Further, number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step 7) Monitoring the evolution of the secondary library

The evolution of the secondary library can be followed as described in Example 1, step j.

Step 8) Identifying molecules of high prevalence

See Example 1, step k

Step 9) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 5

In this Example, a (hypothetical) library composed of beta-peptides is screened for specific interaction of the beta-peptide versus a target molecule. The primary library contains 10^6 beta-peptide tagged X-molecule species. The steps of Example 5 are illustrated in Figures 8A-8B. The two figures should be combined. It is important to note that the screening method used in this Example would apply for other tagged X-molecule species as well. Thus, tagged X-molecule species could have been intrinsic to the X-tag species (one-piece bifunctional tagged X-molecule species) or could have been any chemical entity (d-peptide, gamma-peptide, peptoid, sugar, LNA oligonucleotide, PNA oligomer, small molecule, natural compound, mixed compounds, etc.) with an appended X-tag species (two-piece bifunctional tagged X-molecule species). The steps of Example 5 are illustrated in figure 8 and 9.

Step a) Providing the primary library

The tagged X-molecule species are prepared by performing two alternating parallel syntheses such that a DNA tag species is being chemically linked to the peptide being synthesised (figure 8A). The chemistry for the implementation of this synthesis has been outlined in several publications such as in Nielsen et al. and WO 93/20242.
In this Example, the library (pb1) is built by the combinatorial synthesis of a hexameric peptides formed from 10 different beta-amino acids, which brings the overall diversity of the library to 10^6. Each beta-amino acid is encoded by a particular hexacodon. The employed hexacodons are provided as hexameric phosphoramidites, to reduce the number of couplings in the synthesis of the DNA-tag. However, the hexacodons could also have been formed using six couplings. For the first position of the X-tag, 10 orthogonal codons are used to encode the corresponding beta-aa. For the second position, another 10 orthogonal hexacodons are used and so forth, meaning that a total of 60 orthogonal codons are used, which can be chosen from a total of $4^6 = 4096$ possible hexacodons. The use of orthogonal codons is preferred to reduce faulty hybridisation. (This is particular important for the rate of hybridisation, as it minimizes the time a given X-tag species uses on sampling Y-tag species, before it makes a productive encounter with a 100% complementary Y-tag species.)

At the 3' end, a biotin group is added at a final coupling step during synthesis, to generate tagged X-molecule species as outlined below. The biotin group is added as an affinity handle to facilitate later manipulations of selected tagged X-molecule species. A schematic structure of primary Pb1 (primary beta-peptide molecules is shown in Figure 12. The primary library is used at a concentration of 100 µM in binding buffer.

Step b) Providing the secondary library

The secondary library (pb2) can by synthesised using redundancies as described in Example 1, i.e. that instead of using mono phosporamidites mixtures, hexacodon phosphoramidite mixtures would be used. However, then the coupling efficiencies of individual hexacodon phosphoramidites will have to be further examined to ensure similar coupling efficiency for different hexacodon phoshoramidites.

Instead, the secondary library is prepared in a split-mix combinatorial DNA oligonucleotide synthesis using hexameric anticodons as building blocks, such that each X-tag species will have a complementary counterpart (Y-molecule) in the secondary library. Hexacodon anticodons may also be added using six couplings of mono phosphoramidites.

Fixed regions that enable PCR amplification flank the anti-coding regions of Y-tags. Thus, the Y-molecule species corresponding to the tagged X-molecule species outlined above will be:

5'GATGAT AGTAGT TCGTCG TCAC **CGAGGT TATAGC TAAACC GTTTCG CTCCGA CTAGCT** AGTC ATGATG AGTAGT TGCTGC

Two primers are used for PCR amplification, one of which incorporates a biotin-group into the 5'end of the coding strand of the PCR product:

PCR-primer 1: 5'GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2: 5'bGCAGCA ACTACT CATCAT GACT

The first generation secondary library is used at a concentration of 100 µM.

Step c) Contacting the target molecule with the primary library

The primary library is contacted with the solid phase bound target molecule (e.g. Tumour Necrosis factor alfa) immobilized on sepharose, henceforth also denoted the solid phase. Six µl solid phase (20 µl 30% suspension) is equilibrated in 1000 µl binding buffer 2 (200 mM KCl, 25 mM Tris-HCl, pH 8, 0.01 % Triton X-100) in an eppendorf tube for 5 minutes at 37°C with mixing, whereafter the sample is centrifuged and the binding buffer disposed. This washing procedure is repeated twice to equilibrate the solid phase for incubation with the library. The primary library (100 µl) is then added 100 µl 2xbinding buffer before being incubated with the solid phase at 37°C for 60 minutes with mixing.

Step d) Selecting tagged X-molecule species that interact with the solid phase bound target.

After incubation, the solid phase is washed twice as described above with 1000 µl binding buffer 2 to select tagged X-molecule species interacting with the solid phase.

Step e) Hybridising selected tagged X-molecule species to the secondary library

The secondary library (100µl) is added 1 volume 2xhybridisation buffer, before being added to the solid phase bound target with bound tagged X-molecule species. Next, the sample is heated to 85 °C for 5 minutes, followed by incubation at 65 °C for 12 hours.

Step f) Selecting Y-molecule species hybridised to selected tagged X-molecule species

After hybridisation, 6µl pre-equilibrated streptavidin sepharose is added, and the samples are incubated another 30 minutes at 65°C to immobilize tagged X-molecule species with hybridised Y-molecule species on streptavidin sepharose. Should some tagged X-molecule species have an interaction with the solid phase bound target molecule that is not disrupted during hybridisation, this interaction will serve the same role as immobilisation on streptavidin sepharose (that is to immobilise hybridised Y-molecule species, which allows their selection.) After immobilisation, non-hybridised Y-molecule species are washed away with by two washes with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C.

Step g) Amplifying the selected Y-molecule species

The washed streptavidin sepharose may be used directly as template in the amplification step.Alternatively, hybridised Y-molecule species are eluted with 50 µl 100 mM NaOH using spin filtration, neutralised, ethanol precipitated and dissolved in in 28 µl H_2O of which 25 is aliquoted into 25 standard PCR reactions each containing: 10 µl OptiBuffer, supplied with enzyme, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 1, 2 µl 20 µM PCR-primer 2, 63 µl H_2O and 1 µl BIO-X-ACT™ (4 units) Short DNA polymerase (Bioline Cat. No: BIO-21064). The reaction is cycled 10 times with 94 ° for 30 sec., 55 °C for 30 sec., 68 °C for 60 sec followed by 10 minutes extension at 68 °C. After amplification, all reactions are pooled and the PCR product is gel purified from a 4% agarose gel according

to manufacturers instructions (QIAEX II Gel Extraction Kit, Cat. No. 20021, Qiagen). 400 µl H_2O is used to elute the PCR product from Qiaex II beads.

Step h) Preparation of the next generation secondary library

See Example 1, step h.

Step i) Repetitions

As described in Example 1, step i, the secondary library can be increasingly diluted, because is evolves to contain a larger fraction of Y-molecule species corresponding to active tagged X-molecule species, i.e. if the secondary library is 1000 fold enriched in Y-molecule species corresponding to active tagged X-molecule species, a 1000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have Y-molecule species corresponding to active tagged X-molecule species in moderate excess (5-50 fold) over active tagged X-molecule species for the hybridisation reaction. Further, the number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step j) Monitoring the evolution of the secondary library

The composition of the secondary library is analysed by batch sequencing of the double stranded secondary library. By comparison with the first generation secondary library, it can be determined whether sequence pool is still completely random or whether it has evolved as compared to the starting pool (see also Example 1, step j)

Step k) Identifying molecules of high prevalence

See Example 1, step k.

Step l) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 6

In Example 6, a library composed of 10^9 beta-peptides is screened for activity. Two primary libraries are employed and the secondary library is provided using the alternative method also described in Example 2. Again, it is important to note that the screening method used in this Example would apply for other tagged X-molecule species as well. The steps of Example 6 are illustrated in Figures 9A-9C. The three figures should be combined so that Figure 9A and 9B runs in parallel and continue in Figure 9C.

Step 1) Providing the primary libraries

Tagged X-molecule species are prepared as described in Example 5, except that fixed regions for PCR amplification are added in both ends of the X-tag.

Hexameric beta-peptides are build from 32 monomeric beta amino acids, i.e. the overall diversity of the library becomes $32^6 = 1.1 \times 10^9$. 32 orthogonal hexameric codons are used for each position, i.e. a total of 192 hexameric codons are employed.

Exemplified structures:

Pb2 (primary beta-peptide):
Hexameric betapeptide-GCAGCA ACTACT CATCAT GACT **AGCTAG TCGGAG CGAAAC GGTTTA GCTATA ACCTCG** GTGA CGACGA ACTACT ATCATC-3'

Pb3 (primary beta-peptide):
Hexameric betapeptide-CAGTAG TAGCCA ACGGCT AGTA **AGCTAG TCGGAG CGAAAC GGTTTA GCTATA ACCTCG** ATCG TTAGAC GCTATC CGAGTA-3'

The primary libraries are used at a concentration of 500 µM in binding buffer.
The following PCR primers are used:

PCR-primer 1:	5' GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2:	5'bGCAGCA ACTACT CATCAT GACT
PCR primer-3:	5' TACTCG GATAGC GTCTAA CGAT
PCR primer-4:	5'bCAGTAG TAGCCA ACGGCT AGTA

Second-strand synthesis is performed as described in Example 2.

Step 2) Contacting the target molecule with the primary library

The two primary libraries, pb2 and pb3 are contacted with the solid phase bound target molecule (TNFalfa) in separate experiments, each as described in Example 5, step c

Step 3) Selecting tagged X-molecule species that interact with the solid phase.

After incubation, the solid phase is washed twice with 1000 µl binding buffer to select tagged X-molecule species interacting with the solid phase bound target.

Step 4) Amplifying the selected A-tags

Second strands (anti-coding strands) are eluted from the solid phase with bound tagged X-molecule species, before serving as templates for PCR amplification; the solid phase is resuspended in 60 µl 100 mM NaOH and spinfiltered, whereafter the eluate is neutralised by the addition of 60 µl 100 mM HCl and 15 µl 900 mM Tris-HCl pH 8.5. Subsequently, 126 µl is aliquoted into 63 standard PCR reactions each containing: 10 µl Optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM upstream PCR-primer, 2 µl 20 µM downstream PCR-primer 2, 61 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72°C for 90 sec followed by 10 minutes extension at 72 °C.

For amplification of the pb1 primary library, PCR primers 1 and 2 are employed. Because PCR primer 2 is biotinylated in its 5'end, the resulting PCR product is biotinylated at the 5'end of the coding strand. Likewise, for amplification of pb2, PCR primers 3 and 4 are employed which biotinylates the resulting PCR product at the 5' end of the coding

Step 5) Providing the secondary library

a) The pb1 PCR products and pb2 PCR products are ethanol precipitated and redissolved in 500 µl H_2O. Next, the samples are extracted twice with 200 µl phenol, and one time with

200 µl chloroform followed by immobilization on 100 µl pre-equilibrated streptavidin sepharose.

b) The anti-coding strand of the pb1 PCR product is batch eluted by adding 400 µl 100 mM NaOH to the solid phase followed by centrifugation of the eppendorf tube. After elution, the streptavidin sepharose containing the pb1 coding strand is washed twice with 1000 µl hybridization buffer.

The anti-coding strand of the pb2 PCR product is eluted with 400 µl 100 mM NaOH using spinfiltration. The eluate is subsequently neutralized, whereafter the ssDNA is ethanol precipitated and redissolved in 400 µl binding buffer.

c) The immobilised coding strands of the pb1 PCR product are now hybridised to the complementary anti-coding strands from the pb2 PCR product. Hybridisation is performed by heating the sample to 85 °C for 5 minutes, followed by incubation at 65° for 12 hours.

d) After hybridisation, the streptavidin sepharose is washed two times with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C to select hybridised pb2 anti-coding strands (Y-molecule species)

e) Selected pb2 strands are eluted with 400 µl 100 mM NaOH using spinfiltration, whereafter the eluate is neutralized and ethanol precipitated. The dried precipitate is dissolved in 22 µl H_2O of which 20 µl is aliquoted into 10 PCR reactions each containing: 10 µl optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM PCR-primer 3, 2 µl 20 µM PCR-primer 4, 62 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.

f) The resulting PCR product is immobilized on 15 µl streptavidin sepharose, wherafter the anti-coding strand is eluted with 40 µl 100 mM NaOH, followed by neutralisation and ethanol precipitation. The air dried precipitate is dissolved in 20 µl H_2O to produce the first generation secondary library.

Step 6) Repetitions

In the next round, the pb1 primary library is again selected against the solid phase and selected X-tags PCR amplified and immobilized on streptavidin sepharose. The anti-coding

strands are then eluted and coding strands hybridized to complementary anti-coding Y-molecule species of the first generation secondary library. Hereby selected Y-molecule species are PCR amplified to generate the second-generation secondary library.

As described in Example 1, step i, the secondary library can be increasingly diluted, because is evolves to contain a larger fraction of Y-molecule species corresponding to active tagged X-molecule species, i.e. if the secondary library is 10000 fold enriched in Y-molecule species corresponding to active tagged X-molecule species, a 10.000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have Y-molecule species corresponding to active tagged X-molecule species in moderate excess (5-50 fold) over active tagged X-molecule species for the hybridisation reaction. Further, the number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step 7) Monitoring the evolution of the secondary library

The composition of the secondary library is analysed by batch sequencing of the double stranded secondary library. By comparison with the first generation secondary library, it can be determined whether sequence pool is still completely random or whether it has evolved as compared to the starting pool.

Step 8) Identifying molecules of high prevalence

See Example 1, step k.

Step 9) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 7

Example 7 is an extension of Example 3. Hence, selected tagged X-molecule species are specifically eluted by competition with soluble target molecule. Moreover, a photocleavable

biotin linker is used for manipulation of DNA strands. Again, it is important to note that the screening method used in this Example would apply for other tagged X-molecule species as well. The steps of Example 7 are illustrated in Figures 10A-10C. The three figures should be combined so that Figure 10A and 10B runs in parallel and continue in Figure 10C.

Step 1) Providing the primary libraries

See Example 6

The following PCR primers are used:

PCR-primer 1: 5'GATGAT AGTAGT TCGTCG TCAC
PCR-primer 3: 5' TACTCG GATAGC GTCTAA CGAT
PCR-primer 5: 5'pcb GCAGCA ACTACT CATCAT GACT
PCR-primer 6: 5' CAGTAG TAGCCA ACGGCT AGTA
PCR-primer 7: 5'pcb TACTCG GATAGC GTCTAA CGAT

Step 2) Contacting the primary libraries with the target molecule

The two primary libraries, pb1 and pb2 are contacted with the solid phase bound target (TNFalfa) in separate experiments Example 1, step c

Step 3) Selecting tagged X-molecule species with a specific target molecule interaction

After incubation, the solid phase is washed twice with 1000 µl binding buffer to select tagged X-molecule species interacting with the solid phase bound target. Moreover, tagged X-molecule species bound specifically are eluted using competitive elution; the solid phase is resuspended in 500 µl binding buffer + 1mM soluble target molecule and incubated at 37°C for 5 hours, whereafter the samples are spinfiltered and the liquid phase collected. Subsequently, the the liquid phase is extracted twice with 200 µl phenol, one time with 200 µl chloroform, ethanol precipitated and redissolved in 75 µl binding buffer.

Step 4) Amplifying the selected A-tags

Selected A-tags are PCR amplified as described in Example 3.

Step 5) Providing the secondary library

See Example 3.

Step 6) Repetitions

See Example 6.

Step 7) Monitoring the evolution of the secondary library

See Example 5, step j.

Step 8) Identifying molecules of high prevalence

See Example 1, step k.

Step 9) Identifying tagged X-molecule species with an X-tag (A-tag) species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 8

Example 8 is an extension of Example 6, the only difference being that the hybridisation reaction is performed in solution as also described in Example 4. Again, it is important to note that the screening method used in this Example would apply for other tagged X-molecule species as well. The steps of Example 8 are illustrated in Figures 11A-11C. The three figures should be combined so that Figure 11A and 11B run in parallel and continue in Figure 11C.

Step 1) Providing the primary libraries

See Example 6.

The following PCR primers are used:

PCR-primer 1:	5' GATGAT AGTAGT TCGTCG TCAC
PCR-primer 2:	5'bGCAGCA ACTACT CATCAT GACT
PCR primer-3:	5' TACTCG GATAGC GTCTAA CGAT
PCR primer-4:	5'bCAGTAG TAGCCA ACGGCT AGTA
PCR primer-8:	5' CAGTAG TAGCCA ACGGCT AGTA

The second nucleotide from the 3' end in PCR primer 8 is a ribonucleotide (in bold type).

Second-strand synthesis is performed as described in Example 2.

Step 2) Contacting the target molecule with the primary library

See Example 6.

Step 3) Selecting tagged X-molecule species that interact with the solid phase

See Example 6.

Step 4) Amplifying the selected A-tags

Second strands (anti-coding strands) are eluted from the solid phase with bound tagged X-molecule species, before serving as templates for PCR amplification; the solid phase is resuspended in 60 µl 100 mM NaOH and spinfiltered, whereafter the eluate is neutralised by the addition of 60 µl 100 mM HCl and 15 µl 500 mM Tris-HCl pH 8.5. Subsequently, 126 µl is aliquoted into 63 standard PCR reactions each containing: 10 µl Optibuffer, 16µl 2.5mM dNTP, 6 µl 25 mM $MgCl_2$, 2 µl 20 µM upstream PCR-primer, 2 µl 20 µM downstream PCR-primer 2, 61 µl H_2O and 1 µl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 90 sec followed by 10 minutes extension at 72 °C.

For amplification of the pb1 primary library, PCR primers 1 and 2 are employed. Because PCR primer 2 is biotinylated in its 5'end, the resulting PCR product is biotinylated at the 5'end of the coding strand. Likewise, for amplification of pb2, PCR primers 4 and 8 are employed which biotinylates the resulting PCR product at the 5' end of the coding and introduces a ribonucleotide in the anti-coding strand.

Step 5) Providing the secondary library

a) The pb1 PCR products and pb2 PCR products are ethanol precipitated and redissolved in 500 µl H20. Next, the samples are extracted twice with 200 µl phenol, and one time with 200 µl chloroform followed by immobilization on 100 µl pre-equilibrated streptavidin sepharose.

i) The pb1 PCR product is now immobilized on 100 µl pre-equilibrated streptavidin sepharose and the anti-coding strand of the psn2 PCR product eluted with 400 µl 100 mM NaOH using spinfiltration. Next, the eluate is neutralised, ethanol precipitated and redissolved in 10 µl hybridisation buffer.

ii) The pb2 PCR product is added 1/10 volume 1 M NaOH and incubated at 80 °C for 5 minutes, which cleaves the anti-coding strand at the ribonucleotide reside in PCR primer-10. Next, the sample is neutralised, ethanol precipitated and redissolved in 500 µl formamide loading buffer. The sample is now heated to 94° for 3 minutes and loaded on a 6% denaturing (8 M urea) polyacrylamide gel and the fragments are resolved until the coding strand has reached the middle of the gel. Hereafter, the positions of fragments are determined by UV-shadowing and the gel-piece containing the coding strand is cut out for subsequent passive elution. After elution, the coding strand is ethanol precipitated and redissolved in 10 µl hybridisation buffer.

b) The coding strand of pb1 and the anti-coding strand of pb2 is now mixed for hybridisation in solution. Hybridisation is performed by heating the sample to 85 °C for 5 minutes, followed by incubation at 65° for 12 hours.

c) After hybridisation, the volume is increased to 100 µl by addition of binding buffer, whereafter the sample is added to 6 µl pre-equilibrated streptavidin sepharose and incubated for 30 minutes at 55 °C with mixing.

d) After immobilisation, the streptavidin sepharose is washed two times with 1000 µl 1xhybridisation buffer followed by one wash with wash-buffer (1xSSC+0.01% Triton X-100) buffer for 5 minutes at 65°C to select pb1 strands hybridised to pb2 strands.

e) Selected pb2 strands are eluted with 400 µl 100 mM NaOH using spinfiltration, whereafter the eluate is neutralized and ethanol precipitated. The dried precipitate is dissolved in 22 µl H_2O of which 20 µl is aliquoted into 10 PCR reactions each containing: 10

μl optibuffer, 16μl 2.5mM dNTP, 6 μl 25 mM $MgCl_2$, 2 μl 20 μM PCR-primer 1, 2 μl 20 μM PCR-primer 2, 62 μl H_2O and 1 μl BIO-X-ACT™ Short DNA polymerase (4 units). The reaction is cycled 10 times with 94°C for 30 sec., 55 °C for 30 sec., 72 °C for 60 sec followed by 10 minutes at 72°C.

f) The resulting PCR product is immobilized on 15 μl streptavidin sepharose, wherafter the anti-coding strand is eluted with 40 μl 100 mM NaOH, followed by neutralisation and ethanol precipitation. The airdried precipitate is dissolved in 20 μl H20 to produce the first generation secondary library.

Step 6) Repetitions

In the next round, the pb1 primary library is again selected against the solid phase and selected X-tags PCR amplified. The anti-coding strand from the resulting PCR product is hydrolysed with NaOH and the coding strand purified from PAGE. Purified pb1 coding strands are hybridized to complementary anti-coding Y-molecule species of the first generation secondary library in solution, whereafter hybridised Y-molecule species (pb2 strands) are selected on streptavidin sepharose. Hereby selected Y-molecule species are PCR amplified to generate the second-generation secondary library.

As described in Example 1a, step l, the secondary library can be increasingly diluted, because it evolves to contain a larger fraction of Y-molecule species corresponding to active tagged X-molecule species, i.e. if the secondary library is 10000 fold enriched in Y-molecule species corresponding to active tagged X-molecule species, a 10.000 fold shortage in total amount of the secondary library can be used for hybridisation. The amount of secondary library can also be adjusted to have Y-molecule species corresponding to active tagged X-molecule species in moderate excess (5 –50 fold) over active tagged X-molecule species for the hybridisation reaction. Further, the number of cycles in the PCR reactions can be adjusted in later rounds and carrier nucleic acids may be employed.

Step 7) Monitoring the evolution of the secondary library

See Example 6, step 7

Step 8) Identifying molecules of high prevalence

See Example 1, step k.

Step 9) Identifying tagged X-molecule species with an X-tag species corresponding to the high prevalence Y-molecule species

See Example 1, step l.

Example 9

Two primary libraries were prepared with a diversity of respectively $2,6x10^5$ and $3x10^6$. The two libraries were prepared as outlined in Example 1 and screened in parallel for binding against a solid phase bound target, in this case streptavidin sepharose. The active X-molecule was designed with a photocleavable linker between X-tag and X-molecules. This approach could be used generally to allow specific elution of X-tags corresponding to active X-molecules. Alternatively, the target could be attached to the solid phase by way of a photocleavable linker or active X-molecules could be eluted by competitive elution. The steps of Example 9 are illustrated in Figures 13A-13B. The two figures should be combined.

Step a) Providing the primary library

Two primary libraries were prepared using redundant positions during DNA synthesis as described in Example 1.

Oligonucleotide PL-10e5 has a total diversity of: $2^{18} = 2.6 \times 10^5$. The redundancy of each position is indicated below the sequence.

<u>10^5 primary library preparation:</u>

```
PL-10e5        5'MRKTAA KYMGAG YRYCAC RRYTCT RYRCTC MYKGCA
Redundancy       222111 222111 222111 222111 222111 222111
```

The active oligonucleotide containing a 5'biotin, PS-BamHI to be present in the primary library was synthesised separately with the following sequence

PS-BamHI 5'pbCGCTAT GTTGAC TAGCAG GGATCC ATTCTG ATCGCT

PS-BamHI was diluted into PL-10e5 to create the 10^5 library. The underlined sequence indicates a BamHI restriction site used to monitor evolution of the secondary library.

10^6 primary library preparation:

Oligonucleotide PL-10e6 has a total diversity of: $2^{12} \times 3^6 = 3.0 \times 10^6$. The redundancy of each position is indicated below the sequence.

```
PL(10e6)     5'MRDTAA KYVGAG YRHCAC YMBTGT RYVCTC MYDGCA
Redundancy     223111 223111 223111 223111 223111 223111
```

The active oligonucleotide containing a 5'biotin, PS-NcoI to be present in the primary library was synthesised separately with the following sequence

PS-NcoI 5'pcbCGCTAT GTTGAC TAGCAG CCATGG ATTCTG ATCGCT

PS-NcoI was diluted into PL-10e6 to create the 10^6 library. The underlined sequence indicates an NcoI restriction site used to monitor evolution of the secondary library.

Step b) Providing the secondary library

For each coding DNA oligonucleotide in the primary libraries (tagged X-molecule species), there is a complementary anti-coding DNA oligonucleotide in the secondary libraries (Y-molecule species). Additionally, the secondary library oligonucleotides have fixed regions in both ends to enable PCR amplification.

10^5 secondary library preparation:

SL-10e5:
5'GCCTGTTGTGAGCCTCCTGTCGAATGCMRKGAGYRYAGARYYGTGRYRCTCKRMTTAMYKGCTGAGGTTAT TCTTGTCTCCC

SS-BamHI:
5'GCCTGTTGTGAGCCTCCTGTCGAAAGCGATCAGAATGGATCCCTGCTAGTCAACATAGCGGCTGAGGTTAT TCTTGTCTCCC

SL-10e5 were diluted into SS-BamHI to create the 10^6 secondary library.

10^6 secondary library preparation:

SL-10e6
5'GCCTGTTGTGAGCCTCCTGTCGAATGCHRKGAGBRYACAVKRGTGDYRCTCBRMTTAHYKGCTGAGGTTAT TCTTGTCTCCC

SS-NcoI:
5'GCCTGTTGTGAGCCTCCTGTCGAAAGCGATCAGAATCCATGGCTGCTAGTCAACATAGCG GCTGAGGTTATTCTTGTCTCCC

SL-10e6 were diluted into SS-NcoI to create the 10^6 secondary library.

The sequences in bold are the anti-coding sequences and the flanking sequences are fixed regions for PCR amplification. Again restriction sites are underlined.

PCR primer 11 and PCR primer 12 were used for PCR amplification of the secondary libraries, the latter comprises a biotin and incorporates the biotin-group at the 5'end of the coding strand of the PCR-product, which allows purification of the anti-coding strand, i.e. the next generation secondary library.

PCR-12: GCCTGT TGTGAG CCTCCT GTCGAA

PCR-11: bGGGAG ACAAGA ATAACC TCAGC

Step c-1) Hybridising Y-molecule species of the secondary library with X-tag species of the primary library

DNA oligonucleotides were mixed according to the scheme below to create both the primary and the secondary libraries in a total volume of 90 µl (6xSSC, 0.01% Triton X-100). Two negative controls omitting signal oligonucleotides in secondary libraries were also prepared.

A) Library-10^5:

27 µl 20x SSC
5.4 µl 0.15% Triton X-100
25 µl 200 µM PL-10e5 (MWG, 180304)
25 µl 200 µM SL-10e5 (MWG, 180304)
3.8 µl 5 nM PS-BamHI (DNAtech, 240304)
3.8 µl 5 nM SS-BamHI (DNAtech, 240304)

B) Negative control omitting signal in secondary library, otherwise as A:

27 µl 20x SSC
5.4 µl 0.15% Triton X-100
25 µl 200 µM PL-10e5 (MWG, 180304)
25 µl 200 µM SL-10e5 (MWG, 180304)
3.8 µl 5 nM PS-BamHI (DNAtech, 240304)
3.8 µl 0.01% Triton X-100

C) Library-10[6]:

27 µl 20x SSC
6.4 µl 0.13% Triton X-100
25 µl 200 µM PL-10e6 (MWG, 180304)
25 µl 200 µM SL-10e6 (MWG, 180304)
3.3 µl 0.5 nM PS-NcoI (DNAtech, 240304)
3.3 µl 0.5 nM SS-NcoI (DNAtech, 240304)

D) Negative control omitting signal in primary library, otherwise as C:

27 µl 20x SSC
6.4 µl 0.13% Triton X-100
25 µl 200 µM PL-10e6 (MWG, 180304)
25 µl 200 µM SL-10e6 (MWG, 180304)
3.3 µl 0.5 nM PS-NcoI (DNAtech, 240304)
3.3 µl 0.01% Triton X-100

Next, the libraries (samples A to D) were heated to 94 °C for 5 minutes followed by incubation at 65 °C ON (18h).

Step d-1) Contacting the target molecule with at least a subset of the primary library hybridised to the secondary library

100 µl solid phase bound target suspension (30% Streptavidin Sepharose High Performance beads in 20% EtOH, Amersham, 17-5113-01) was centrifuged to pellet the solid phase. The supernatant was disposed and 600 µl 6xSSC, 0.01% Triton X-100 added. After resuspension of the solid phase, it was again pelleted by centrifugation and the supernatant disposed. The solid phase was then resuspended in 600 µl 6xSSC, 2 µg/µl

tRNA (170 µl 7 µg/µl tRNA (tRNA from Roche, 109 541, phenol extracted;)+ 180 µl 20x SSC + 250 µl H_2O), centrifugated and the supernatant disposed. Finally, the equilibrated solid phase was resuspended in 70 µl 6xSSC, 0.01% Triton X-100 to give a total volume of app. 100 µl. 20 µl equilibrated solid phase was added to samples A-D from step e. The samples were then incubated at 65 °C for 20 minutes with mixing in a table shaker to allow interaction between the primary library and the solid phase.

Step e-1) Selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule, thereby also selecting Y-tags hybridised to selected X-tags

After incubation with the solid-phase, the samples were transferred to spin-off filters (Ultrafree-MC filter microporous 0.22 micron, Millipore, UFC3 0GV NB) and centrifuged at 3000 rpm for 2x 1 minute. In the first wash, samples were added 300 µl 10xwash buffer (1 M NaCl, 100 mM Tris-HCl pH 8)+0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute. For the second wash, the samples were added 300 µl 1xwash buffer+ 0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute.

Step f-1) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

The solid phase from above with selected X and Y-molecules might be used directly in PCR, but to minimize amplification of non-hybridised Y-tags, X-molecules with hybridised Y-tags were photocleaved of the solid phase, by way of a photocleavable linker between the X-molecules and its corresponding X-tag. The solid phase was resuspended in 100 µl 1xwash buffer, 0.01% Triton X-100 and placed on a UV table for 3 minutes. The released complexes (Y-tags hybridised to X-molecules) were collected by centrifugation at 3000 rpm for 2x 1 minute.

4 PCR mixes were prepared each containing:
308 µl H_2O
55 µl 10xbuffer (Bioline, BIO-21050)
16.5 µl 50 mM $MgCl_2$ (Bioline, BIO-21050)
22 µl dNTPs, 5 mM each (Bioline, BIO-39025)
22 µl 10 µM PCR-11
22 µl 10 µM PCR-12

5.5 µl polymerase (BIO-X-ACT long, Bioline, BIO-21050)

As negative controls, 41 µl was collected from each of the above PCR mix and each added 9 µl 1xwash buffer, 0.01% Triton X-100.

The remaining 410 µl of the PCR mixes was added 90 µl of the samples **A-D** from step f) and each aliquoted in 100 µl in 5 PCR tubes.

Amplification was performed according to the following program:

Initial denaturation:	94 °C, 5 min.
30 cycles:	94°C, 30 sec
	58°C, 60 sec
	72°C, 10 sec.
Final extension:	72°C, 5 min

After amplification, identical PCR samples were pooled.

Step j) Monitoring the evolution of the secondary library

5 µl of negative control samples A-D were added 5 µl H_2O and 2.5 µl of a 25 bp DNA ladder (Promega, #G4511) was added 7.5 µl H_2O. The samples were added 3 µl 30% glycerol and resolved on a 4% GTG (BioWhittaker (BMA), 50084) agarose gel using 1xTBE as running buffer. As expected, no PCR products had formed (data not shown).

5 µl of samples A and B were added 1 µl BamHI + 1 µl 10x BamHI buffer + 1 µl 10x BSA + 2 µl H_2O; 5 µl of samples C and D were added 1 µl NcoI + 1 µl buffer 4 (NEB, B7004S) + 3 µl H_2O.

For comparison, samples with 1 µl H_2O instead of restriction enzyme were also prepared. All were incubated at 37 °C for 2 hours and then added 3 µl 30% glycerol and resolved on a 4% GTG agarose gel. As can be seen in Figure 14, no digestion was seen for any of the samples. Thus, the experiment was continued with another round.

Step h) Preparation of the next generation secondary library

Only the anticoding strand of the PCR product from above is desired and was therefore purified. 100 µl 30% Streptavidin Sepharose High Performance beads in 20% EtOH were centrifuged, the supernatant disposed and 600 µl 6x SSC, 0.01% Triton X-100 added. After resuspension of the streptavidin sepharose, it was again pelleted by centrifugation and the supernatant disposed. The streptavidin sepharose was then resuspended in 70 µl 6x SSC, 0.01% Triton X-100 to give a total volume of app. 100 µl.

The app 480 µl sample A-D from step g were added 200 µl 20x SSC + 20 µl of the above equilibrated streptavidin sepharose. Next, samples A-D were incubated at RT for 20 minutes with mixing. Then the samples were transferred to spin-off filters (2x 370 µl) and centrifuged at 3000 rpm for 2x 1 minute. In the first wash, samples were added 300 µl 10xwash buffer + 0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute. For second wash, the samples were added 300 µl 1xwash buffer + 0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute. Next, samples A-D were resuspended in 40 µl 100 mM NaOH by pipetting up and down a few times and then incubated at RT for 5 minutes. The anticoding strands were then collected by centrifugation at 13000 rpm for 1 minute. 40 µl of the eluted samples were neutralised by adding 40 µl 100 mM HCl + 18 µl 1 M Tris pH 8 + 2 µl 0.5% Triton X-100. Next, the samples were desalted by gel-filtration on G25 columns (MicroSpin G-25 columns, Amersham, 27-5325-01). Finally, 2 µl of the purified samples A-D together with 1, 2 and 4 pmol of the SL-10e6 oligo and 2.5 µl of the 25 bp DNA ladder were analysed on a 4% GTG agarose gel. From the gel, the concentration the purified samples A-D were estimated to be around 1 µM ready for the next round of screening (data not shown).

Step l) Repetitions - Second round

Step c-1) Hybridising Y-molecule species of the secondary library with X-tag species of the primary library

Primary libraries were prepared again and mixed with the second generation secondary libraries from the previous step h) and aliquoted into tubes A-D according to the scheme below to give a total volume of 90 µl (6xSSC, 0.01% Triton X-100).

Three different concentrations of the secondary libraries were used.

A-1) Library-10e5:

27 µl 20x SSC

9 µl 0.1% Triton X-100
25 µl 200 µM PL-10e5
25 µl Sample A
3.8 µl 5 nM PS-BamHI

A-2 and A-3 were as A-1, except that 10-fold and 100-fold diluted Sample A was used.

B-1) Negative control omitting signal in secondary library, otherwise as A:

27 µl 20x SSC
9 µl 0.1% Triton X-100
25 µl 200 µM PL-10e5
25 µl Sample B
3.8 µl 5 nM PS-BamHI

B-2 and B-3 were as B-1, except that 10-fold and 100-fold diluted Sample B was used.

C-1) Library-10e6:

27 µl 20x SSC
9 µl 0.1% Triton X-100
25 µl 200 µM PL-10e6
25 µl Sample C
3.3 µl 0.5 nM PS-NcoI

C-2 and C-3 were as C-1, except that 10-fold and 100-fold diluted Sample C was used.

D-1) Negative control omitting signal in primary library, otherwise as C:

27 µl 20x SSC
9 µl 0.1% Triton X-100
25 µl 200 µM PL-10e6
25 µl Sample D
3.3 µl 0.5 nM PS-NcoI

D-2 and D-3 were as D-1, except that 10-fold and 100-fold diluted Sample D was used.

Next, the libraries (samples A-1 to D-3) were heated to 94 °C for 5 minutes followed by incubation at 65 °C ON (18 h).

Step d-1) Contacting the target molecule with at least a subset of the primary library hybridised to the secondary library

300 µl solid phase bound target suspension (30% Streptavidin Sepharose High Performance beads) was centrifuged to pellet the solid phase. The supernatant was disposed and 1800 µl 6xSSC, 0.01% Triton X-100 added. After resuspension of the solid phase, it was again pelleted by centrifugation and the supernatant disposed. The solid phase was resuspended in 1800 µl 6xSSC + 2 µg/µl tRNA and after centrifugation the supernatant disposed. Finally, the solid phase was resuspended in 210 µl 6xSSC, 0.01% Triton X-100 to give a total volume of app. 300 µl. 20 µl equilibrated solid phase from above was added to sample A-1 to D-3. The samples were incubated at 65 °C for 20 minutes with mixing in a table shaker. The samples were then incubated at 65 °C for 20 minutes with mixing in a table shaker to allow interaction between the primary library and the solid phase.

Step e-1) Selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule, thereby also selecting Y-tags hybridised to selected X-tags

Next the samples were transferred to spin-off filters and centrifuged at 3000 rpm for 2x 1 minute. In the first wash, samples were added 300 µl 10xwash buffer (1 M NaCl, 100 mM Tris-HCl pH 8) + 0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute. For second wash, the samples were added 300 µl 1xwash buffer + 0.01% Triton X-100 and centrifuged at 3000 rpm for 2x 1 minute.

Step f-1) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

The solid phase from above might be used directly in PCR, but to enhance selection of hybridised Y-molecules, X-molecules with hybridised Y-tags were photocleaved of the solid phase: The solid phase were resuspended in 100 µl 1xwash buffer, 0.01% Triton X-100 and placed on the UV table for 3 minutes. The released complexes (Y-tags hybridised to X-molecules) were collected by centrifugation at 3000 rpm for 2x 1 minute.

One PCR mix was prepared containing:
420 µl H_2O
75 µl 10xbuffer (Bioline, BIO-21050)

22.5 µl 50 mM $MgCl_2$ (Bioline, BIO-21050)
30 µl dNTPs, 5 mM each (Bioline, BIO-39025)
30 µl 10 µM PCR-11
30 µl 10 µM PCR-12
7.5 µl polymerase (BIO-X-ACT long, Bioline, BIO-21050)

The mix was aliquoted to 13x 41 µl in PCR tubes and 9 µl of Samples A-1 to D-3 and 1xwash buffer + 0.01% Triton X-100 (negative control) added.

Amplification was performed according to the following program:

Initial denaturation:	94 °C, 5 min.	
30 cycles:	94°C, 30 sec	
	68°C, 60 sec	
	72°C, 10 sec.	
Final extension:	72°C, 5 min	

Step j) Monitoring the evolution of the secondary library

5 µl of samples A-1 to B-3 and the negative control was added 1 µl BamHI + 1 µl 10x BamHI buffer + 1 µl 10x BSA + 2 µl H_2O. Further 5 µl of samples C-1 to D-3 and the negative control were added 1 µl NcoI + 1 µl buffer 4 + 3 µl H_2O.
For comparison, samples with 1 µl H_2O instead of restriction enzyme were also prepared. All were incubated at 37 °C for 2 hours and then added 3 µl 30% glycerol and resolved on a 4% GTG agarose gel.

Figure 15 shows +/- restriction enzyme of sample A-1 to B-2 and 25 bp DNA ladder (2.5 µl).
Figure 16 shows +/- restriction enzyme of sample B-3, neg. PCR Control (BamHI), C-1 to C-3 and 25 bp DNA ladder (2.5 µl).
Figure 17 shows +/- restriction enzyme of sample D-1 to D-3, neg. PCR Control (NcoI) and 25 bp DNA ladder (2.5 µl).

Results and conclusion

Approximately 30% of sample A1 and about 80 % of samples A-2 and A-3 could be restricted by BamHI. This means that the secondary library had evolved from containing 1 SS-BamHI oligonucleotides per 260.000 library oligonucleotides into containing between 30 and 80 SS-BamHI oligonucleotides per 100 library oligonucleotides. This reflects an enrichment of app. 80.000 (for A1) and 210.000 fold (for A2 and A3).

Approximately 5% of sample C1 and about 20% of sample C-2 and C-3 could be restricted by NcoI. This means that the secondary library had evolved from containing 1 SS-NcoI oligonucleotides per 3.000.000 library oligonucleotides into containing between 5 and 20 SS-NcoI oligonucleotides per 100 library oligonucleotides. This reflects an enrichment of app. 130.000 (for C1) and 520.000 fold (for C2 and C3). Importantly, no restriction was seen in any of the controls.

It can therefore be concluded that the present invention successfully has been used to detect binders in a non-evolvable primary library (comprising non-amplifiable molecules) by the use of a secondary evolvable library.

REFERENCES

Abelson: Methods in Enzymology, Academic Press, vol. 267, edited by John Abelson.

Anderson: Anderson, Nucleic Acid Hybridisation, Springer Verlag, 1998

Ausubel et al: Short Protocols in Molecular Biology, Ausubel et al, John Wiley and Sons, 1992

Houghten et al.: Houghten RA, Pinilla C, Blondelle SE, Appel JR, Dooley CT, Cuervo JH. Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. Nature 1991 Nov 7;354(6348):84-6

Nielsen et al.: Nielsen, J and Janda, KD: Toward chemical implementation of encoded combinatorial libraries, Methods, 1994, 361-371

Lam et al.: Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ. A new type of synthetic peptide library for identifying ligand-binding activity. Nature 1991 Nov 7;354(6348):82-4

Liu et al.: Gartner ZJ, Liu DR. The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. J Am Chem Soc 2001 Jul 18;123(28):6961-3

Needels et al.: Needels MC, Jones DG, Tate EH, Heinkel GL, Kochersperger LM, Dower WJ, Barrett RW, Gallop MA Generation and screening of an oligonucleotide-encoded synthetic peptide library. Proc Natl Acad Sci U S A 1993 Nov 15;90(22):10700-4.

Olejnik et al 1: Olejnik J, Krzymanska-Olejnik E, Rothschild KJ. Photocleavable affinity tags for isolation and detection of biomolecules. Methods Enzymol 1998;291:135-54.

Olejnik et al 2: Olejnik J, Krzymanska-Olejnik E, Rothschild KJ. Photocleavable aminotag phosphoramidites for 5'-termini DNA/RNA labelling. Nucleic Acids Res 1998 Aug 1;26(15):3572-6.

Sambrook et al: Molecular cloning: a Laboratory Manual: 3nd edition, Volume 1 and 2, Sambrook et al., 2001, Cold Spring Harbor Laboratory Press

US 5,635,400

WO 02/103008, Gouliaev et al.

WO 93/20242, Lerner et al.

CLAIMS

1. A method of selecting, among a plurality of molecules, a molecule that is capable of specifically interacting with a target molecule, the method comprising the steps of

a) providing a secondary library comprising a plurality of Y-molecule species, each Y-molecule species comprising a specific tag species (Y-tag species),

b) providing a primary library comprising a plurality of tagged X-molecule species, wherein the tagged X-molecule species comprises an X-molecule species and a specific tag species (X-tag species), and wherein at least one X-tag species of the primary library is capable of hybridising to at least one Y-tag species of the secondary library,

c) contacting the target molecule with at least a subset of the primary library,

d) selecting the tagged X-molecule species of the primary library that interact specifically with the target molecule,

e) optionally, contacting the secondary library with the X-tag species of the selected tagged X-molecule species,

f) selecting Y-molecule species from the secondary library that are capable of hybridising with an X-tag species of a selected tagged X-molecule species of step d) or are capable of hybridising with the complementary sequence of an X-tag species of a selected tagged X-molecule species of step d),

g) amplifying the selected Y-molecule species, the product of the amplification process being a secondary library,

h) repeating steps a) , f) and g), wherein the secondary library provided in step a) is derived from a secondary library produced in a previous step g),

i) identifying Y-molecule species of high prevalence in a generation of the secondary library, and

j) identifying, from the primary library, X-molecule species corresponding to the Y-tag species of the Y-molecule species of high prevalence.

2. The method according to claim 1, wherein the number of repetitions of step a), f) and g), as described in step h), is at least 2 times, such as at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20, 30 or such as at least 40 times.

3. The method according to any of the preceding claims, wherein the primary library provided in step b) is substantially identical in every repetition.

4. The method according to claims 1 or 2, wherein the primary library provided in step b) is different from the initial primary library in at least one of the repetitions.

5. The method according to any of the preceding claims, further comprising, in at least one repetition, a step of monitoring the amplification product of step g).

6. The method according to claim 5, wherein the result of the monitoring is used for determining if a new repetition of step a), f) and g) should be performed.

7. The method according to any of the preceding claims, wherein the primary library comprises at least 10^3 tagged X-molecule species, such as at least 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, such as at least 10^{15}.

8. The method according to any of the preceding claims, wherein the concentration of a tagged X-molecule species is at least 10^{-18} M.

9. The method according to any of the preceding claims, wherein the concentration of a tagged X-molecule species at most 1 mM.

10. The method according to any of the preceding claims, wherein the primary library further comprises an aqueous solvent.

11. The method according to any of the preceding claims, wherein the primary library further comprises an organic solvent.

12. The method according to any of the preceding claims, wherein, the primary library further comprising an additive selected from the group consisting of a detergent, a preservative, a pH buffer and a salts.

13. The method according to any of the preceding claims, wherein the secondary library comprises at least 10^1 Y-molecule species, such as at least 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, such as at least 10^{15}.

14. The method according to any of the preceding claims, wherein the lowest concentration of a Y-molecule species is 10^{-22} M.

15. The method according to any of the preceding claims, wherein the highest concentration of a Y-molecule species is 1 mM.

16. The method according to any of the preceding claims, wherein the secondary library of step a) is derived from X-tag species of selected tagged X-molecule species of a previous step d).

17. The method according to claim 16, wherein a secondary library of step a) is provided by a method comprising the following steps

1) providing a library comprising a plurality of tagged X-molecule species, wherein the tagged X-molecule species is provided with an amplifiable tag species (A-tag species), said A-tag species comprises a tag species and at least one primer binding site for amplifying said tag species,

the tagged X-molecule species are characterised by being divided into two sub-libraries of tagged X_1-molecule species and tagged X_2-molecule species, wherein the amplifyable tag species (A_1-tag) of the X_1-molecule species is different from the amplifyable tag species (A_2-tag) of the X_2-molecule species,

2) contacting a target molecule with the sub-library of tagged X_1-molecule species,

3) selecting, from the sub-library of tagged X_1-molecule species, tagged X_1-molecule species that interact specifically with the target molecule,

4) contacting a target molecule with the sub-library of tagged X_2-molecule species,

5) selecting, from the sub-library of tagged X_2-molecule species, tagged X_2-molecule species that interact specifically with the target molecule,

6) amplifying the A_1-tag species from the selected tagged X_1-molecule species by hybridizing specific primers to the primer binding site of the A_1-tag species, and performing the amplification thereby obtaining the anti-coding parts of the selected A_1-tag species,

7) amplifying the A_2-tag species from the selected tagged X_2-molecule species by hybridising specific primers to the primer binding site of the A_2-tag species, and performing the amplification thereby obtaining the anti-coding parts of the selected A_2-tag species,

8) selecting the coding part of the selected A_1-tag species and selecting the anti-coding part of the selected A_2-tag species,

9) contacting the coding part of the selected A_1-tag species with the anti-coding part of the selected A_2-tag species under conditions that allow for stringent hybridisation,

10) selecting the anti-coding A_2-tag species of step 9) that hybridise to selected coding A_1-tag species, and

11) using the selected anti-coding A_2-tag species of step 10) as secondary library.

18. The method according to 17, wherein where step 11) of claim 17 further comprises at least one step selected from the groups of steps consisting of

11a) amplifying the selected anti-coding A_2-tag species,

11b) purifying the amplification product, and

11c) diluting the amplification product.

19. The method according to any of the preceding claims, wherein the tagged X-molecule species comprises an X-tag species linked to an X-molecule species, said X-tag species comprising a tag species.

20. The method according to any of the preceding claims, wherein the X-tag species is linked to the X-molecule species via a linker molecule or via a direct binding.

21. The method according to claim 20, wherein a bond involved in direct binding or in the linking using a linker molecule is of a covalent character or of a non-covalent character.

22. The method according to claim 20 or 21, wherein the linker molecules is selected from the group consisting of a di-aldehyde such as a glutaraldehyde, a polymer such as a oligosacharide (oligedextran), a nucleic, and a peptide.

23. The method according to any of the claims 20-22, wherein the linker molecule comprises at least two active groups, said active groups are capable of further polymerisation.

24. The method according to any of the claims 20-23, wherein the polymer of the linker molecule comprises at least 2 monomers such as at least 5, 10, 15, 20, 50, 100 such as at least 200 monomers.

25. The method according to any of the claims 20-24, wherein the polymer of the linker molecule is at least 5 Å long such as at least 10 Å, 15 Å, 20 Å, 30 Å, 50 Å, such as at least 1000 Å long.

26. The method according to any of the claims 20-25, wherein the polymer of the linker molecule is substantially linear.

27. The method according to any of the claims 20-25, wherein the polymer of the linker molecule is substantially unbranched or branched.

28. The method according to any of the preceding claims, wherein the tagged X-molecule species further comprises a capture component.

29. The method according to 28, wherein the capture component is selected from the group consisting of a biotin, an avidin, a streptavidin, an antibody and functional derivatives thereof.

30. The method according to any of the preceding claims, wherein the tagged X-molecule further comprises a release component.

31. The method according to 30, wherein the release component is located in the X-molecule, or between the X-molecule and the linker molecule, or in the linking molecule, or between the linker molecule and the X-tag species, or in the X-tag species, or between the capture component and the X-tag species.

32. The method according to claims 30 or 31, wherein the release component is selected from the group consisting of a selective cleavage site for an enzyme, a cleavage site for a nucleic acid restriction enzyme, a ribonucleotide, a photocleavable group.

33. The method according to 32, wherein the photocleavable group is a o-nitrobenzyl linker.

34. The method according to any of the preceding claims, wherein the tagged X-molecule species is prepared using a method comprising the steps of

a) providing a linker molecule comprising at least a first functional group and a second functional group, said first functional group is capable of receiving a tag codon group, said second functional groups is capable of receiving an X-group

b) adding a new tag codon group to the first functional group, said new tag codon group being capable of receiving a further tag codon group,

c) adding a new X-group to the second functional group, said new X-group being capable of receiving a further X-group.

35. The method according to claim 34, wherein step b) and c) is performed in the same reaction mixture.

36. The method according to claims 34 or 35, wherein the X-group comprises at least one component selected from the group consisting of an amino acid, a nucleotide, a carbohydrate, a carbohydrate, derivatives thereof and any combinations thereof.

37. The method according to claim 36, wherein the amino acid is selected from the group consisting of an alanine, an arginine, an asparagine, an aspartic acid, a cysteine, a glutamine, a glutamic acid, a glycine, a histidine, an isoleucine, a leucine, a lysine, a methionine, a phenylalanine, a proline, a serine, a threonine, a tryptophan, a tyrosine, a valine and a synthetic amino acid.

38. The method according to any of the preceding claims, wherein the X-molecule species comprises a component selected from a group consisting of an a peptide, a nucleic acid, a protein, a receptor, a receptor analogue, a polysaccharide, a drug, a hormone, a hormone analogue and an enzyme.

Figure 1A

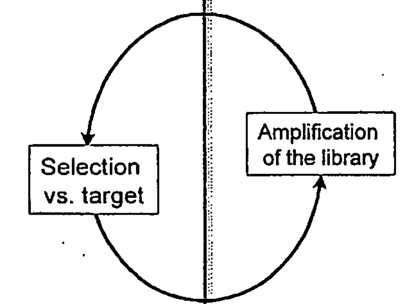


Figure 1B

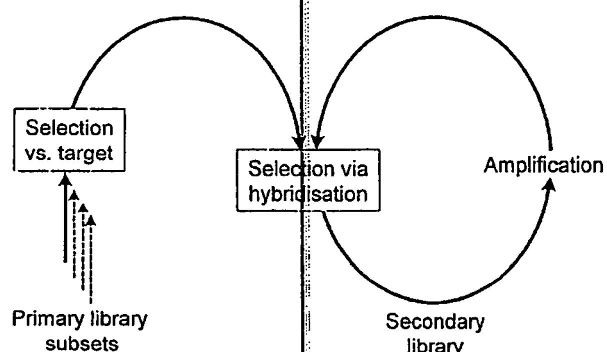


Figure 2A

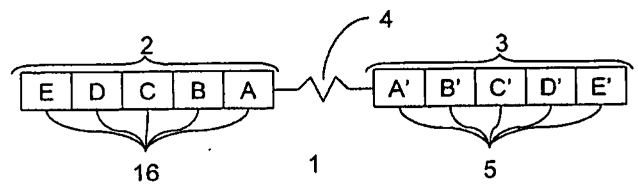


Figure 2B

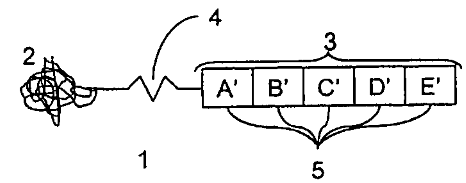


Figure 2c

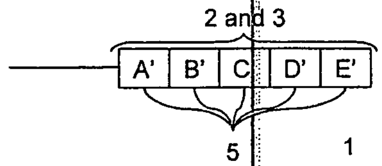


Figure 2D

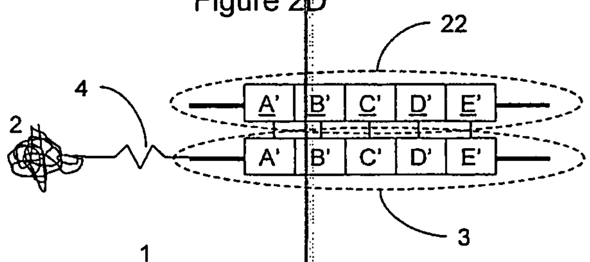


Figure 3A

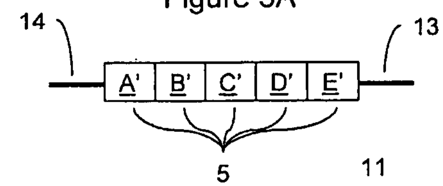


Figure 3B

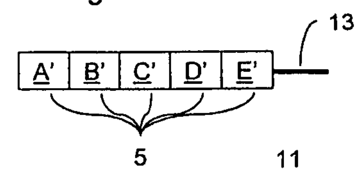


Figure 4A

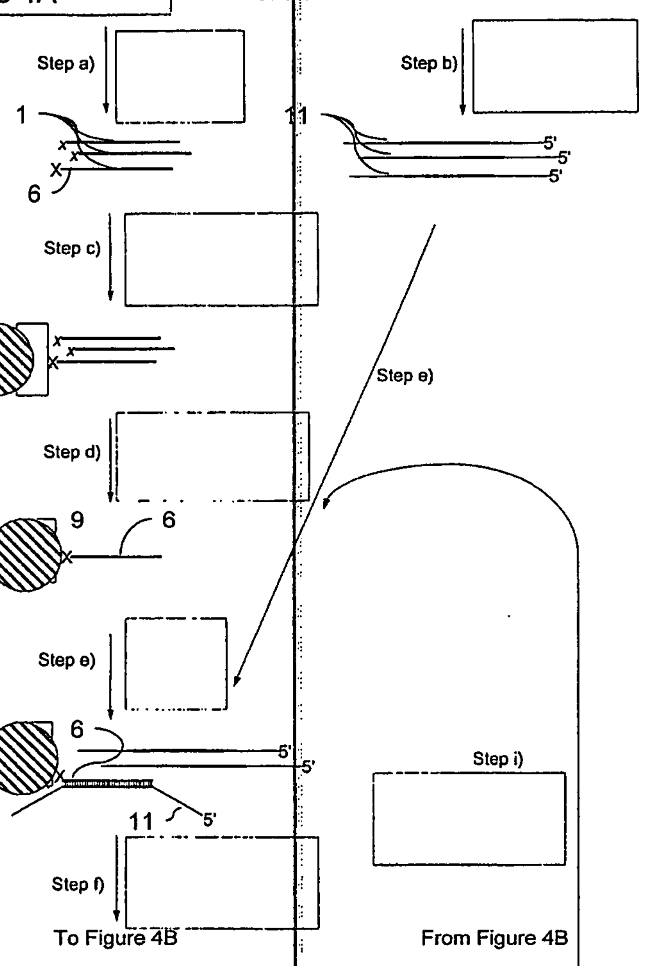


Figure 4B

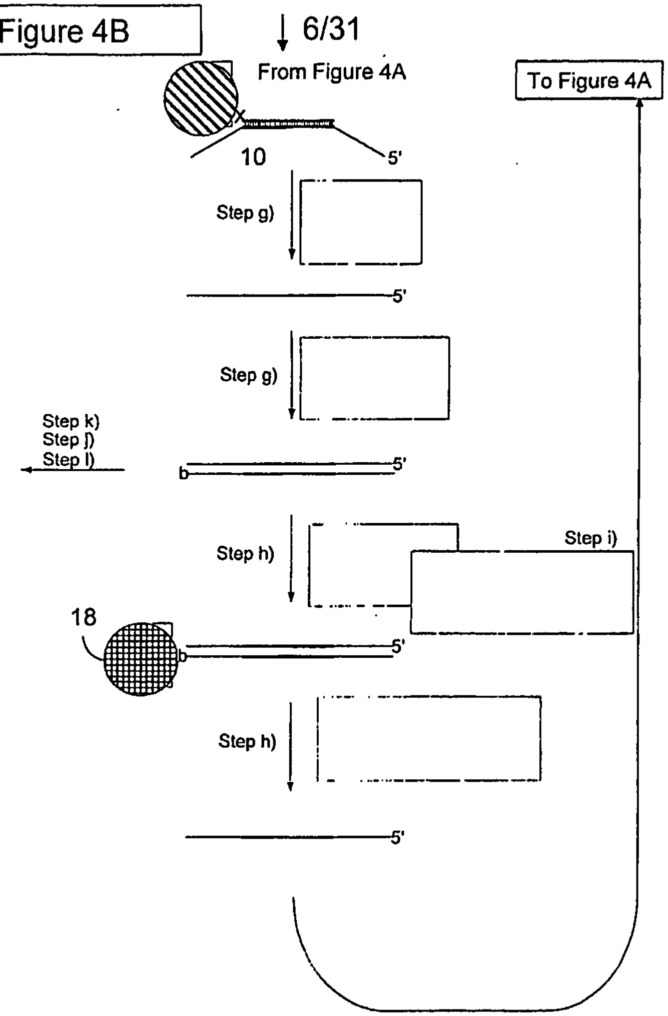
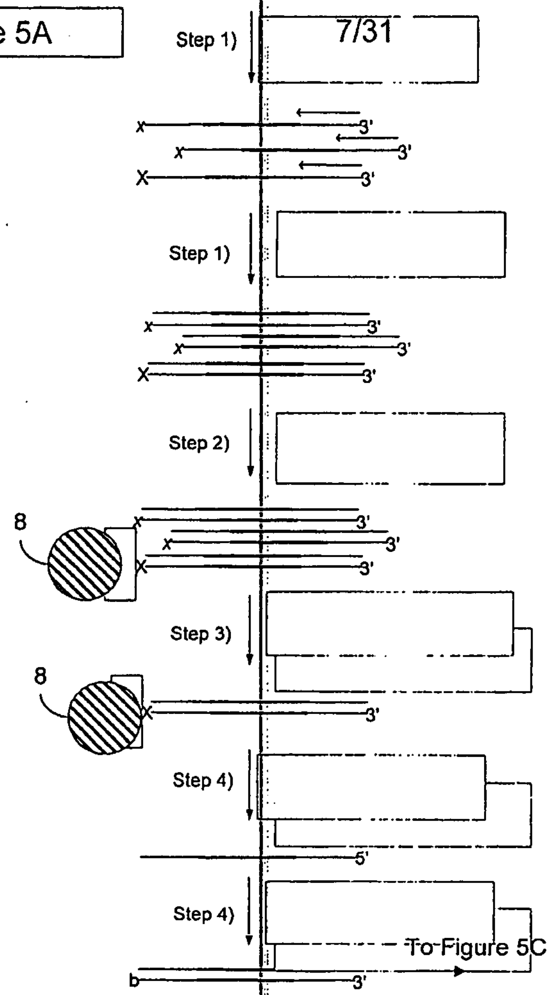


Figure 5A



WO 2004/099441

PCT/DK2004/000325

Figure 5B

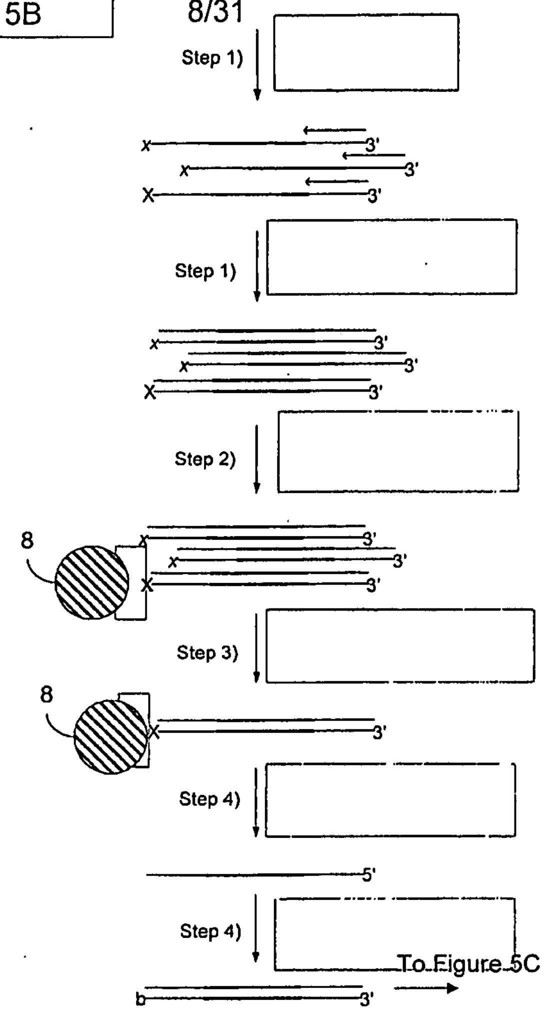


Figure 5C

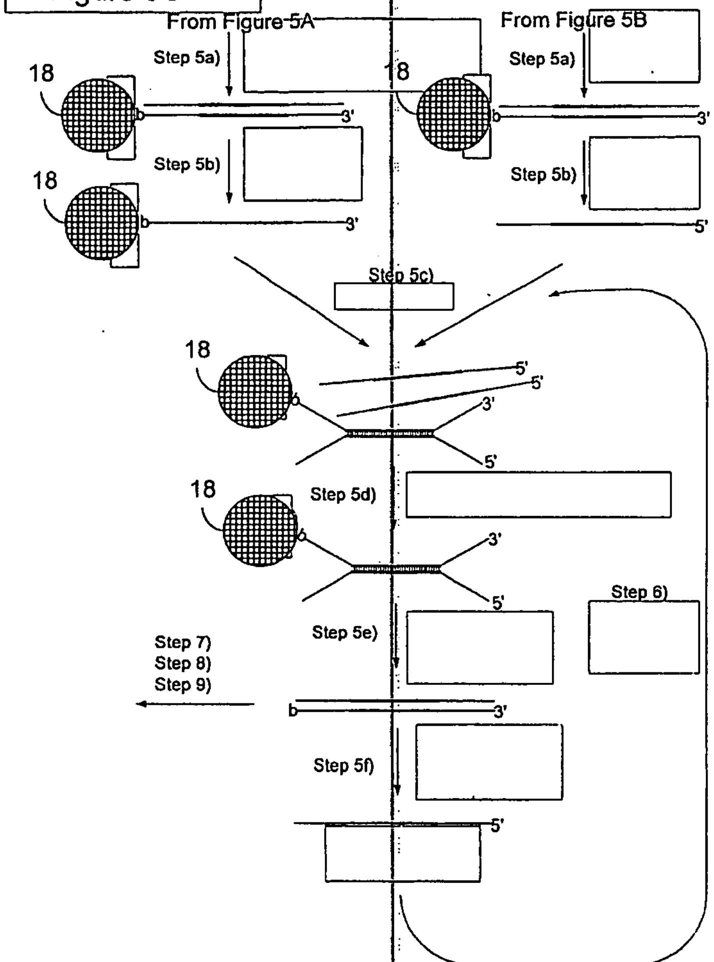


Figure 6A

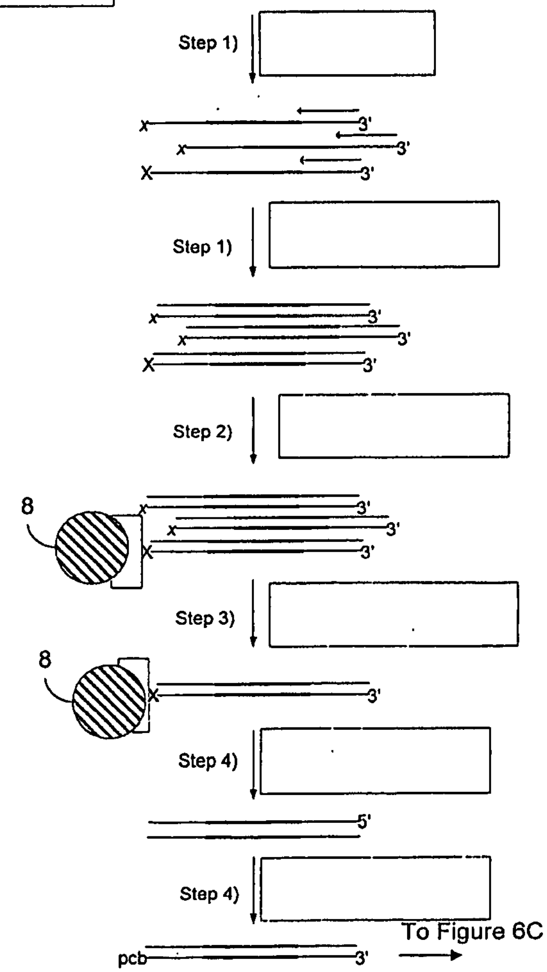


Figure 6B

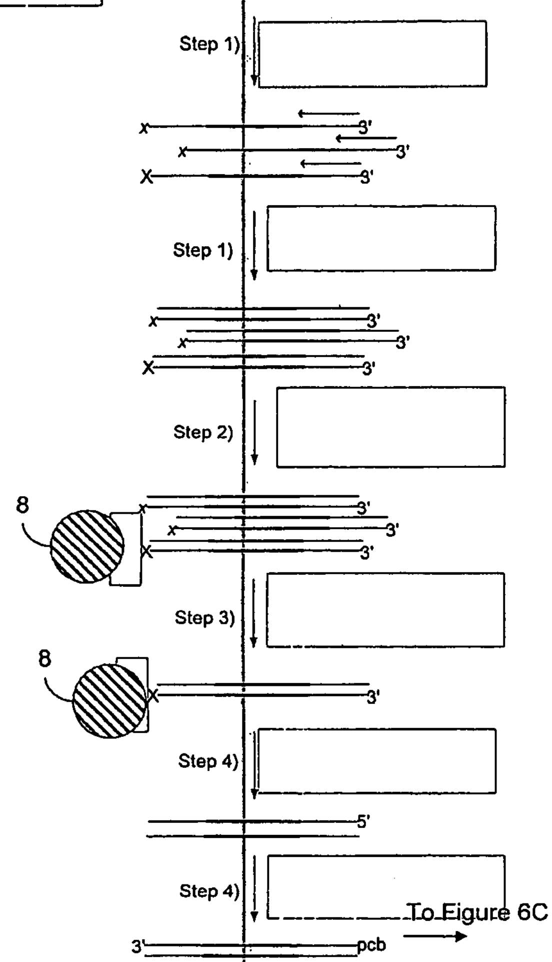


Figure 6C

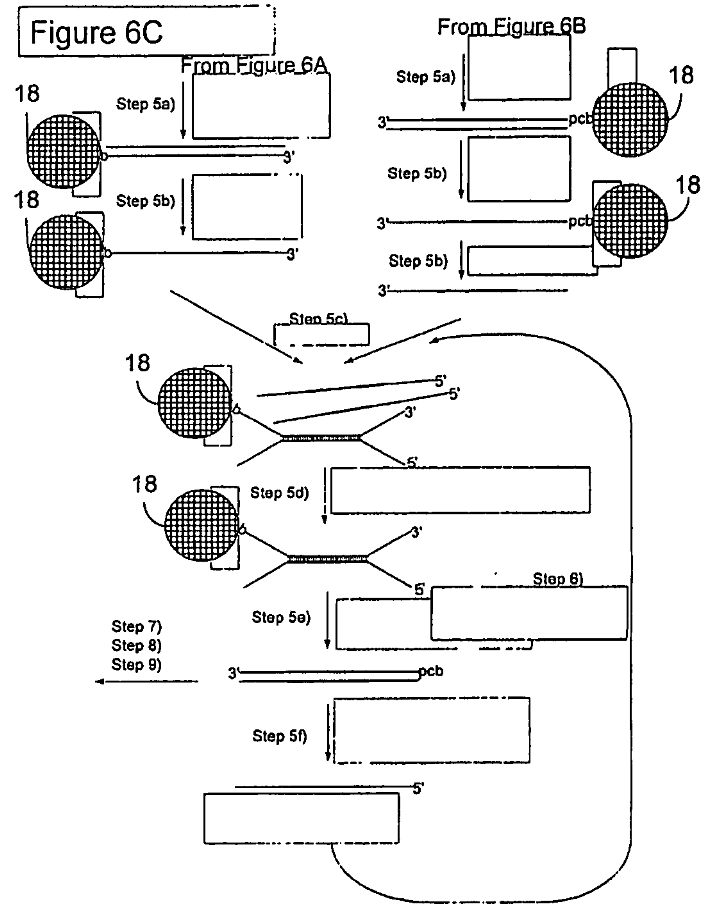


Figure 7A

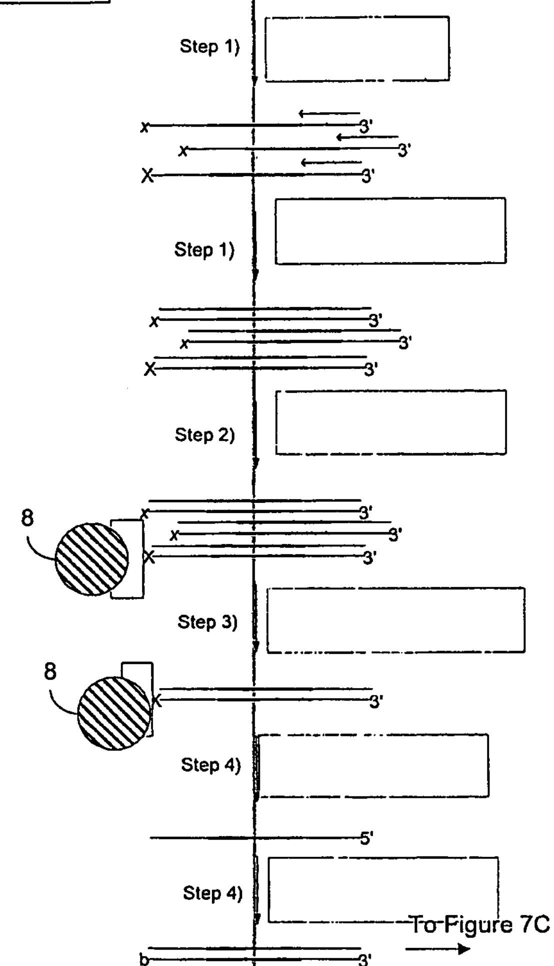


Figure 7B

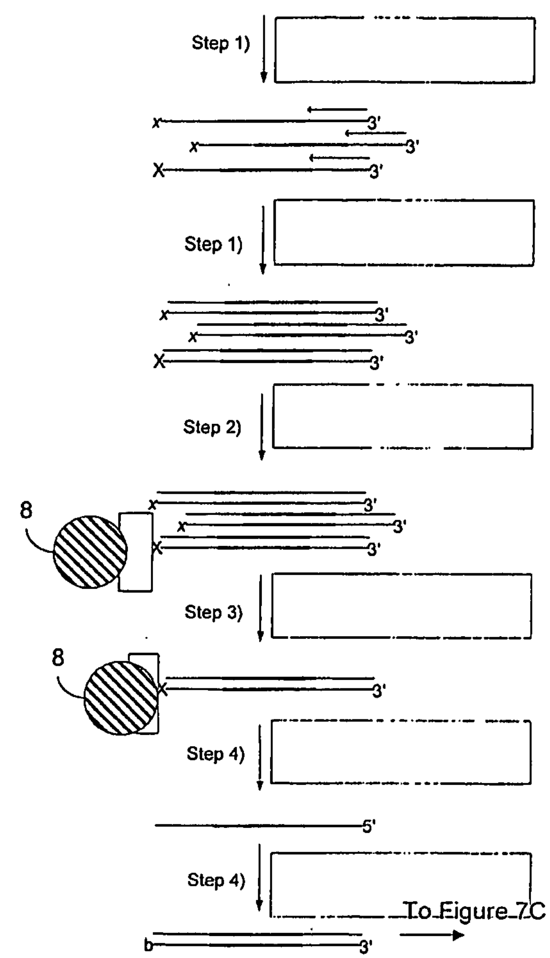
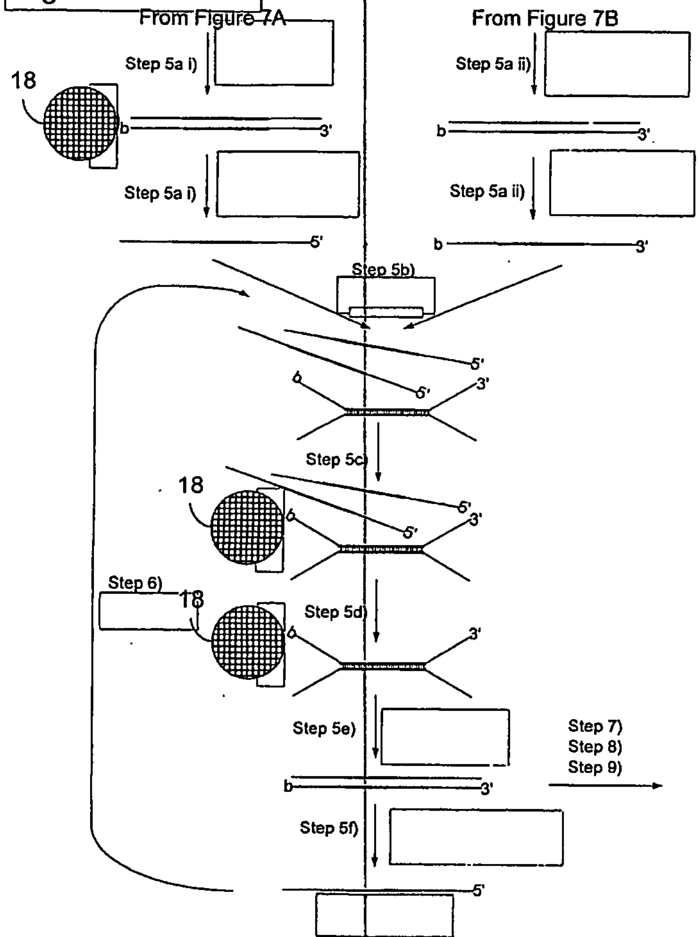


Figure 7C



WO 2004/099441
PCT/DK2004/000325

Figure 8A

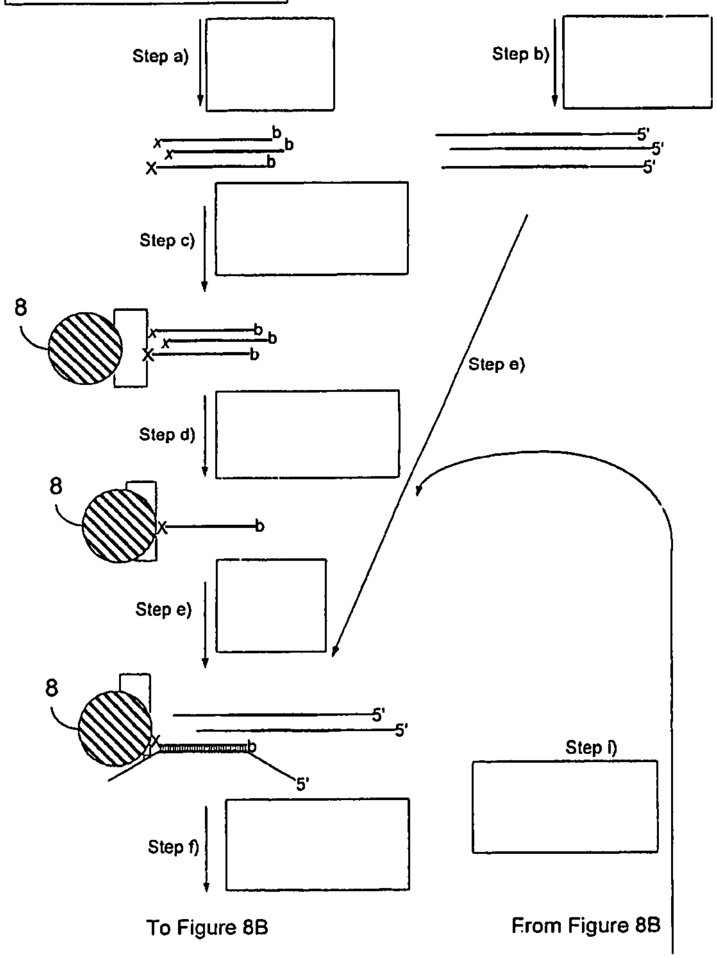


Figure 8B

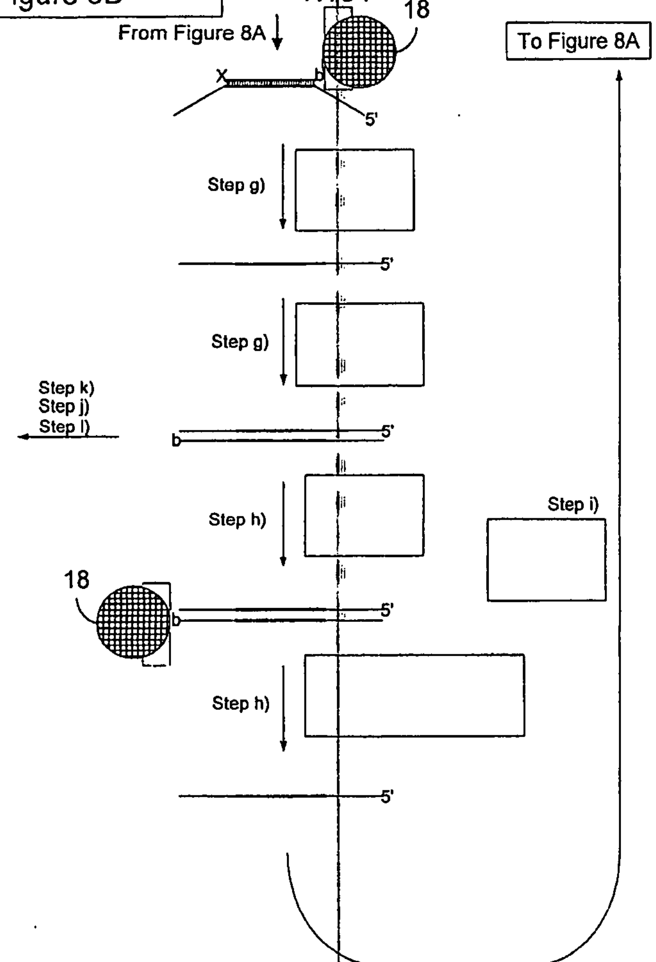


Figure 9A

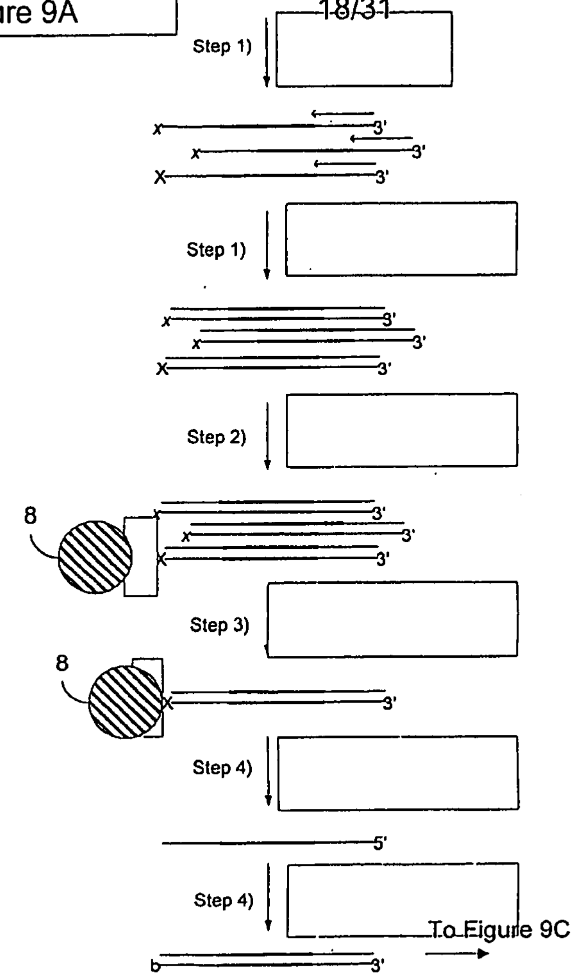


Figure 9B

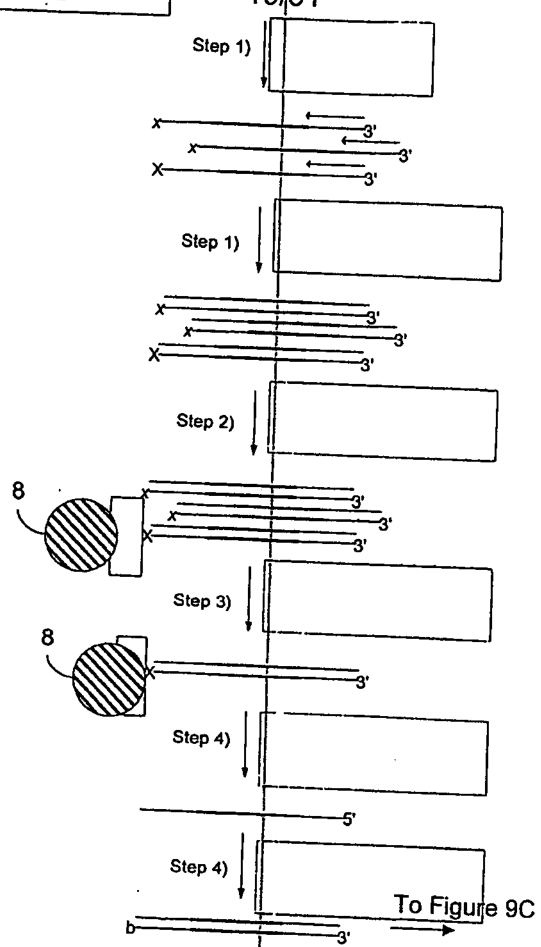


Figure 9C

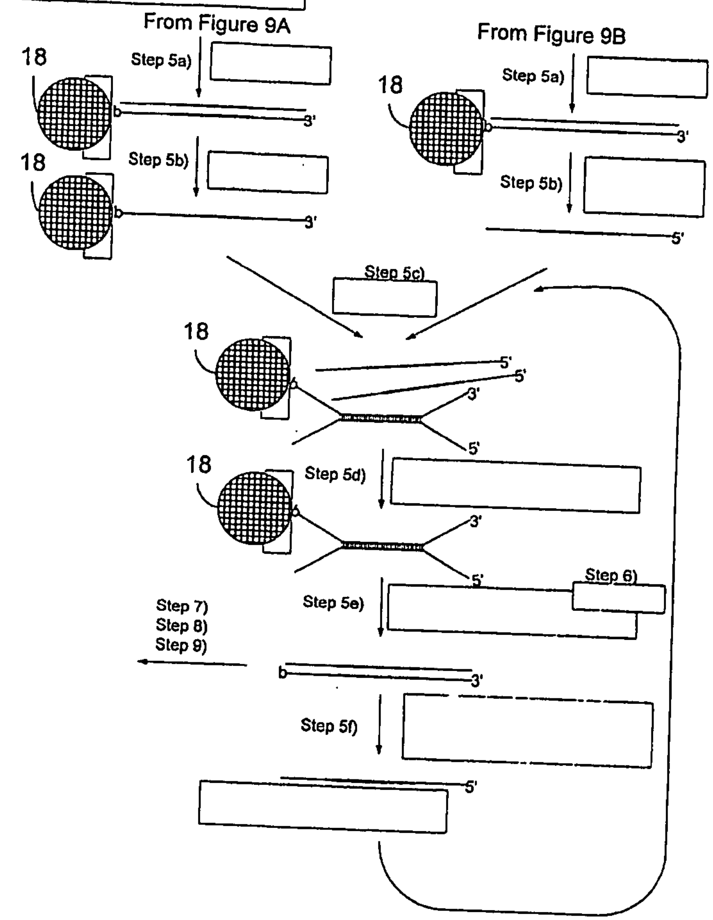
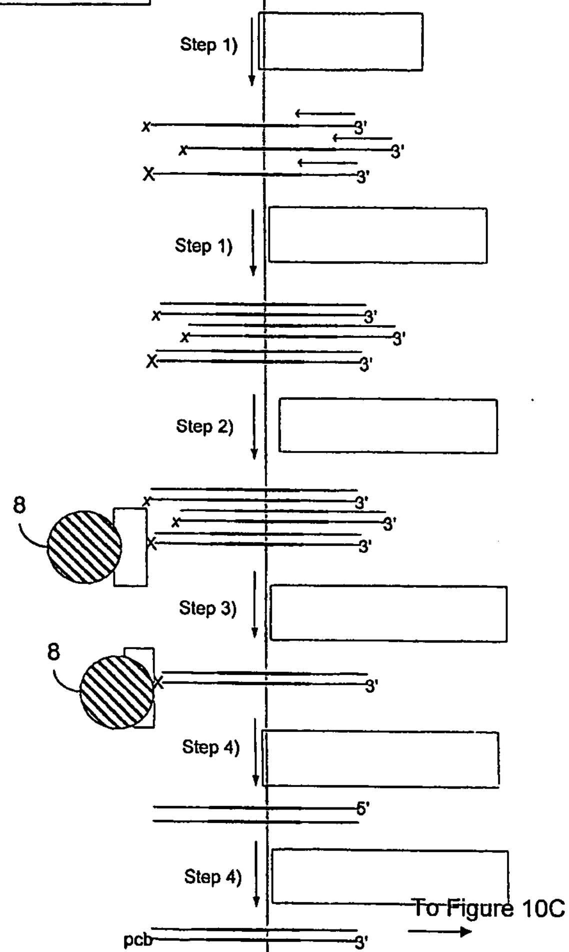


Figure 10A



WO 2004/099441

PCT/DK2004/000325

Figure 10B

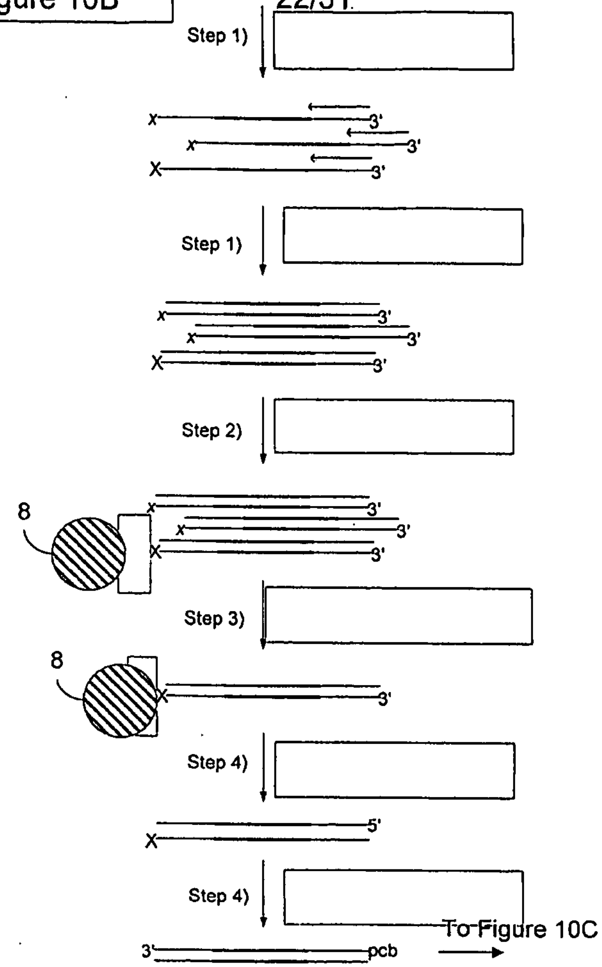
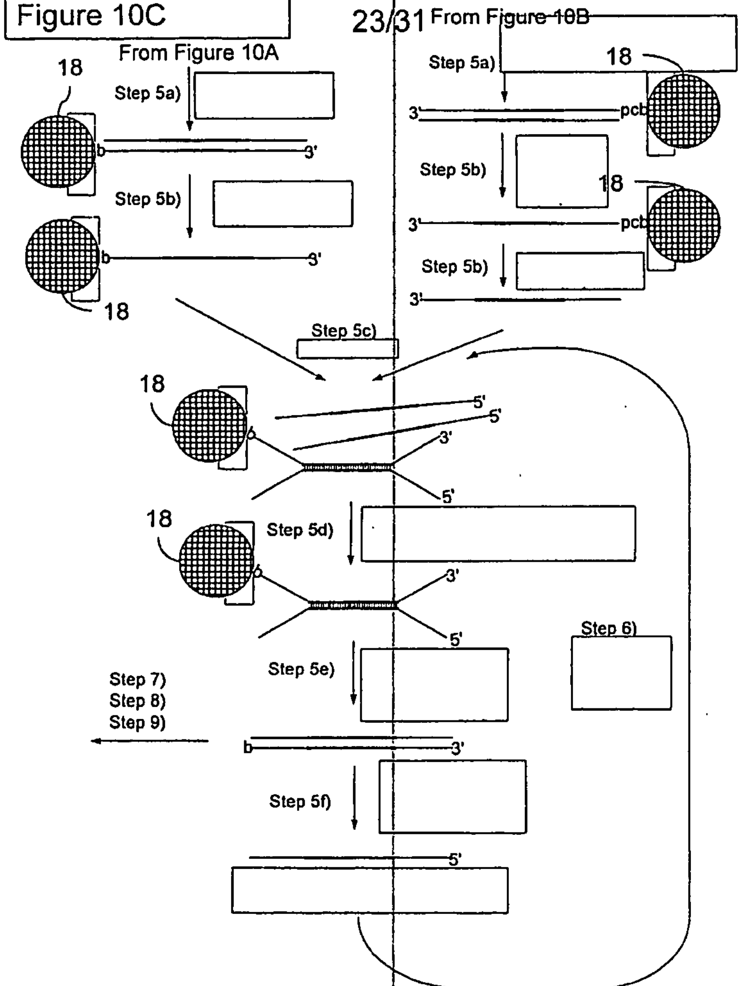


Figure 10C



WO 2004/099441

PCT/DK2004/000325

Figure 11A

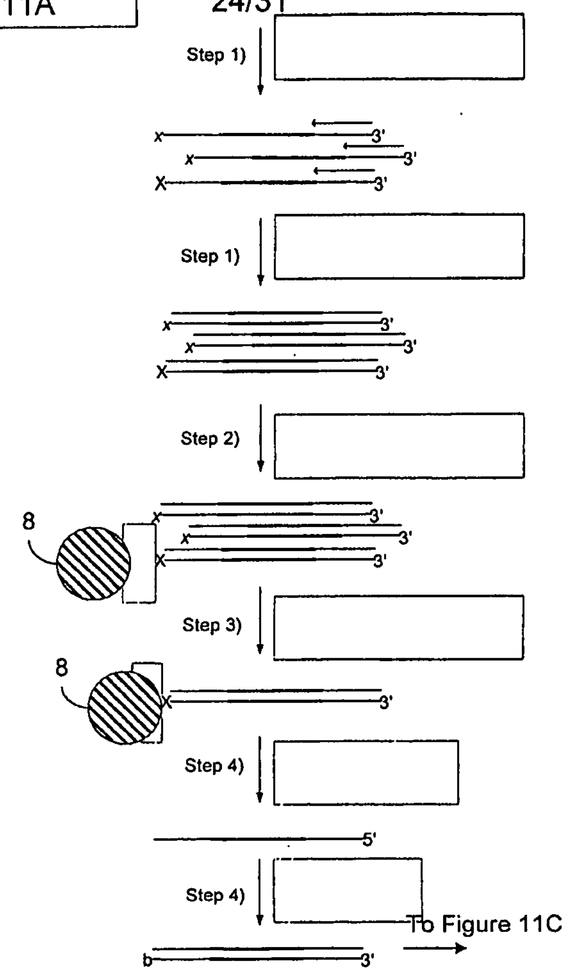


Figure 11B

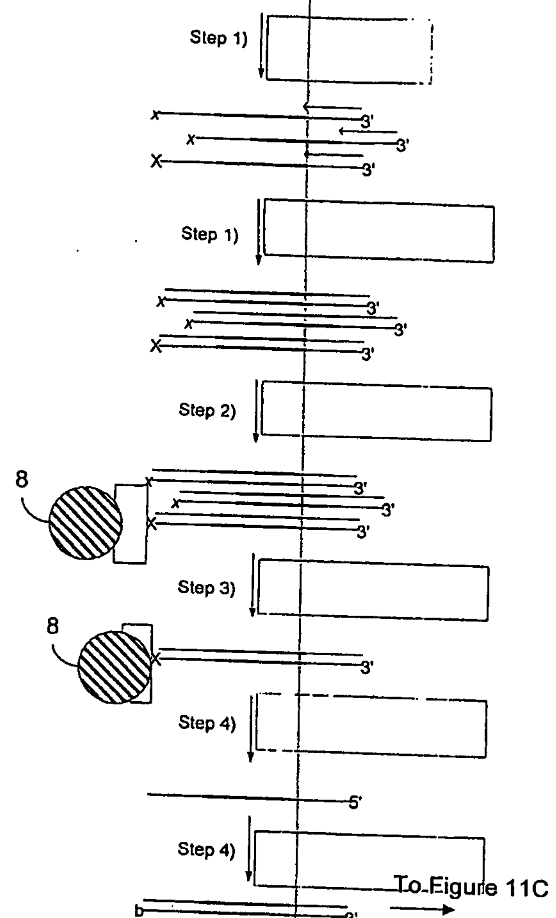


Figure 11C

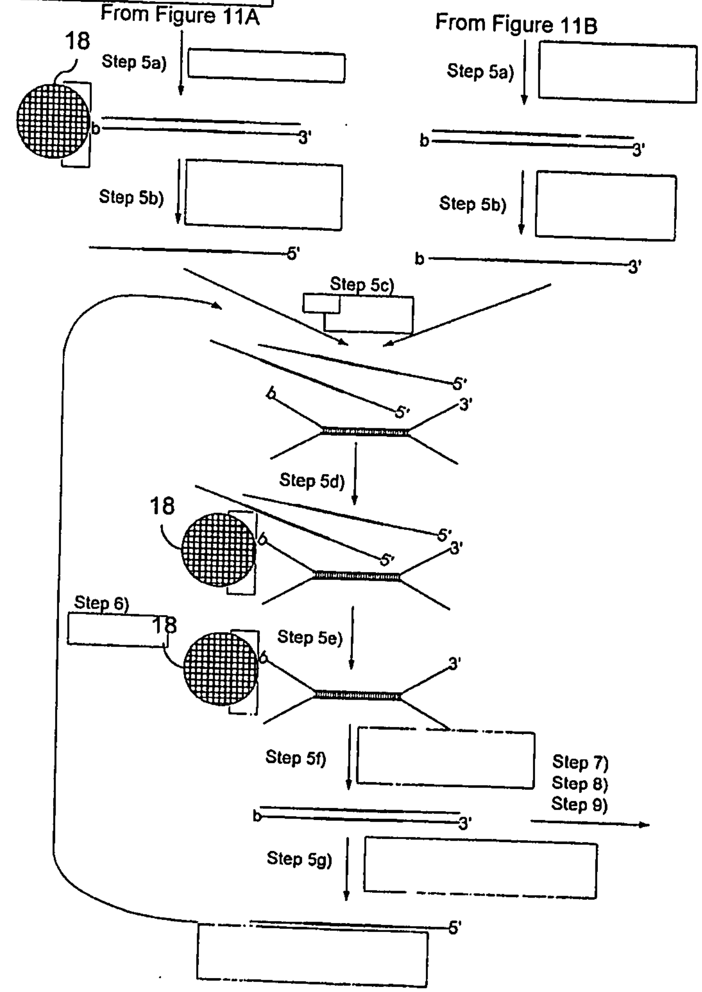


Figure 12

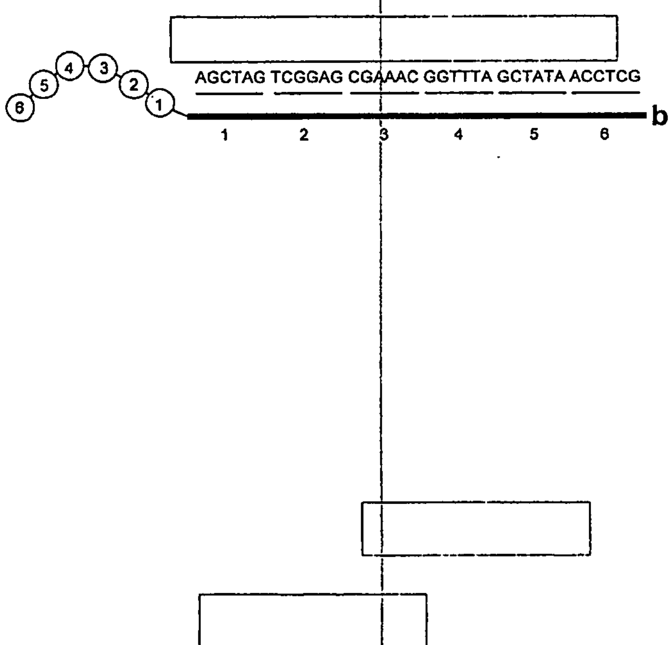


Figure 13A

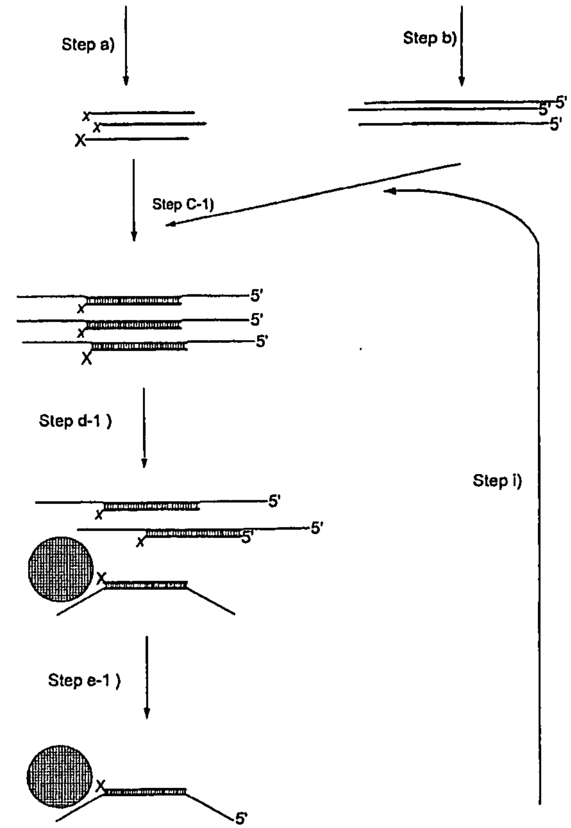


Figure 13B

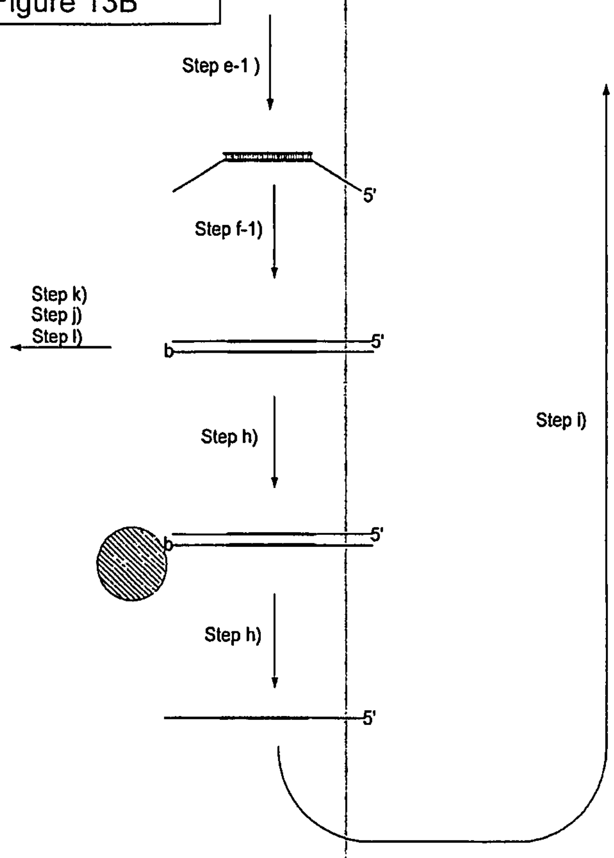


Figure 14

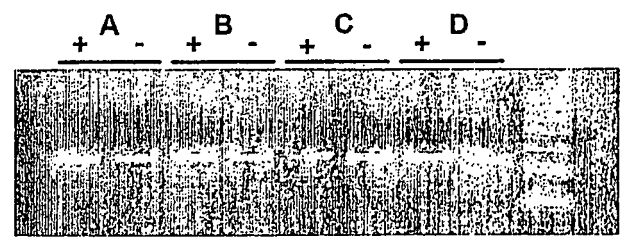


Figure 15

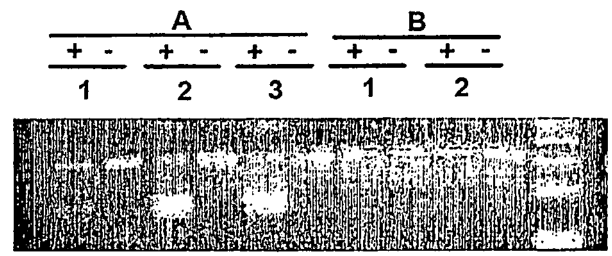


Figure 16

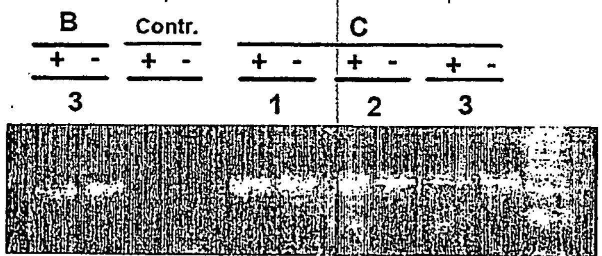


Figure 17